

FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTTCAGATCTGCTCGGTAGA
CCTGGTGCACCACCACC**ATG**TTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG
GTTTTCCACCCAGCTTTTCACCAAGGCCTCCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA
ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA
GAACTGGCCAAGAAGCTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT
GATCAGATGGGAAGATGGTTTGTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTTGTGCTA
CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT
ATGTCAAGGATAGAATTTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAACAGCT
TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACCTTCATGATGAGAGGCTCTTG
GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC
CATATGACAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT
GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTTCTCATCAGAGCTGCATGGTACAC
AGCTGGCATTGTGGGAGGCCCTCTCCACTGTGGCCATGTGTGCGCCAGTGAAAAGTTTCTGA
ACATGGGTGCACCCCTGGGAGTGGGCCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG
TTTCTTCCACCTACCACCGTGGCTGGTGGCACTCTTACTCAGTGGCAATGTACGGTGGATT
AGTTCTTTTCAGCATGTTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT
CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT
ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAA**TG**
AAGTGAAGTCAAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAATGGGGCAGATATGC
ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA
TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT
AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT
TTGGTGAATGTGAAAATAAAGTTTGTGTGATGAGAATGTAAGTCTTTTTTCTACTTTAAA
TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT
GCAGAAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGGACCAGAGAGAAG
GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA
GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTCAATTT
GCTGAACCTTAACAAAAGTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG
CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC
ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA
ATAACTTTAAAAGTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG
AATACAAACAGTATACTCATG

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWFVAGGAAVGLGALCYYGGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLITILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFMSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
21

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCCGTCTCCGCCTTCTGCAT
 CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGGCCGCGTCTGAGGG
 GGTCCGCACGGGGAGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGTTCGAAG**ATG**TCGG
 ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCCGCCACCGTC
 GCCGTGCCCTTGGTCCGCAAACTCGGCCTCATCAGCCCGGCCTACCTCTTCCTCTGGCCCGA
 AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG
 GTCCAGGAACTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA
 CTTGAAACAGGAGCTTTTGTATGGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTTAACTG
 GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCTCTTGATCA
 TGTCAGTACTTTATGTCTGGGCCAGCTGAACAGAGACATGATTGTATCATTTTGGTTTGGGA
 ACACGATTTAAGGCCTGCTATTTACCCTGGGTTATCCTTGGATTCAACTATATCATCGGAGG
 CTCGGTAATCAATGAGCTTATTGGAAATCTGGTGGACATCTTTATTTTTTCTTAATGTTCA
 GATACCAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGCTGG
 CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC
 TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC
 AG**TGA**AGGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCTCCAGTGCTGGGTG
 CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC
 AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCACAAGTTTCACGAT
 TCTCATTTCAAGTCCCTTACTGCTGTGAAGAACAATACCAACTGTGCAAATTGCAAACTGAC
 TACATTTTTTTGGTGTCTTCTCTTCTCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCTT
 AATCTGCTGGCATTTGAGCTGGGGCTGGGTACCAAACCTTCCCAAAGGACCTTATCTCTT
 TCTTGCACACATGCCTCTCTCCCACTTTTCCCAACCCCAACATTTGCAACTAGAAAAAGTTG
 CCCATAAAATTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGTC
 ACAACAATCATATTCACGTTATTTTCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAG
 AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATGCTTCGACTGACATCCGTTGTT
 AACCCTTTGCCACTCTTCAGATATTTTTTATAAAAAAGTACCACTGAGTTCATGAGGGCCA
 CAGATTGGTTATTAATGAGATACGAGGGTTGGTGTGGGTGTTTGTTCCTGAGCTAAGTGA
 TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCCC
 TTTGCGTTTCATATGTAGCCCTACTGGCTTTGGTGTAGCTGGAGTAGTTGGGTTGCTTGTGT
 TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG
 ATTCCCATTTCAATCTCATTTCTGGATATGTGTTCAATTGAGTAAAGGAGGAGAGACCCTCATA
 CGCTATTTAAATGTCACTTTTTTGCCTATCCCCCGTTTTTTTGGTTCATGTTTCAATTAATTGT
 GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTAAAGCTAATGTAAGCACATCTA
 AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA
 TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTTT
 TCGTAGGTGGGCTTTTCTTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA
 TGGCTTTTACACAGTTATTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT
 ATTGAGTGGCTGTACACTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTGATCAGTTTCTT
 TTCAGGAACATTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT
 GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG
 TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA
 TGTTTTGTTGTTTCACTGTGGCCACAATAAAGTTTACTTGTAAATTTTAGAGGCCATTACT
 CCAATTATGTTGCACGTACACTCATTTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA
 TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCCCTCTTACCAGTCAGCTGCCTGCGAG
 CAGTCATTTTTTCTTAAAGGTTTACAAGTATTTAGAAGTTTTCAGTTCAGGGCAAATGTTT
 ATGAAGTTATTCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT
 ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTGTACTTGACAAGGCAAACA
 TGACAGTGGATTCTCTTTACAAATGGAAAAAAATCCTTATTTTGTATAAAGGACTTCCC
 TTTTGTAAACTAATCCTTTTTTATTGGTAAAAATGTAAATTAAATGTGCAACTTG

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGGLSPAYLFLWPEAFLYRFQIWRPITATFYF
PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLEMLLFNWICIVITGLAMDMQLLMIP
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIIGGSVINELIGNLVGHLYFFL
MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL
GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG
 GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT
 CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC
 CCTTTGGGGCGGGATGGCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG
 CGGGGTTCCCTGCGAGGCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT
 GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT
 GGCCTGTGTTCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC
 ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT
 AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC
 CATTTTGCACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA
 AAAACATTGAAATGCAGCTGCAAGCCATTGCAATAATTCAAGAGAGAAATGGTGTATTACCT
 GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT
 GAGGGAAGTTCTTAGAAAATCAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAGGAAAA
 AACAGTTATCAGAGGCTAAAACAGAAGAGCCACAGTGCATTCCAGTGAAGCTGCAATAATG
 AATAATTCCCAAGGGGATGGTGAACATTTTGCACACCCACCCCTCAGAAGTTAAAATGCATTT
 TGCTAATCAGTCAATAGAACCTTTGGGAAGAAAAGTGGAAGGTCTGAAACTTCCTCCCTCC
 CACAAAAGGCCTGAAGATTCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC
 TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA
 TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG
 GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG
 CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAAACTCAAAGAAGAAGTTATTAATAAGTA
ATAATTAAGAACAATTTAACAAAATGGAAGTTCAAATTGTCTTAAAAATAAATTATTTAGTC
 CTTACACTG

FIGURE 6

MAAEEEEDEVWVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP
LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVVSLEHEEMKILREVL
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEEKQTLLKRRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTTCGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAGAGGAATATGACCAGGAA

FIGURE 8

GCGTGGTTTTTGTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTG
 TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG
 TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTA AAAAAC
 AGTGGAATGGAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCACAATGTATAC
 ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT
 CTGCCAATGAAGAAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTG
 GTGAAGCTAGTTTTCTGTGTGCTTGTGTCAATTCTGTGTTATAAAGAAAGATCATCAAAGTAG
 AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCTGCCT
 TTCTTTATTTCTGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATG
 GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTAGGATAGTGCTGAA
 GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCT
 TGAAGTGGCGGGACTAAACTTTACAGCACAACCTTGGCAGGACGTGGATTTTCATCACGATGCC
 TTTTTCAGCCCTTCCAATTCCTGCCTTCTTTTCAGAAAGTGAAGTGTCCAGAAAAGACAATTG
 TACAGCAAAGGAATGGACTTTTCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTCAGTC
 ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTGTTTTATTTCTTCAATGGCT
 AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA
 GAACAGCAAACCTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA
 GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTATGGCCACAGTGCATTTTCAGTAGCC
 CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA
 CATGTTCCATGTCTTGATGGCCCAGGTTACCACTGTCAATTATCACAAACAGTGTCTGTCTGCTG
 TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGAAGCCCCATCAGTCCTTCTCTCTATA
 TTTATTTATAATGCCAGCAAGCCTCAAGTTCGGAATACGCACCTAGGCAAGAAAGGATCCG
 AGATCTAAGTGGCAATCTTTGGGAGCGTTCAGTGGGGATGGAGAAGAACTAGAAAGACTTA
 CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTCTAACTGGTACCCACATAGTTTTGCA
 GCTCTCTTGAACCTTATTTTTCACATTTTCAGTGTGTTGTAATATTTATCTTTTCACTTTGATA
 AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT
 CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA
 ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT
 TTTCTTGGCCTTCAAGCTTCCAAAAAAGCTTGTAAATAATCATGTTAGCTATAGCTTGTATAT
 ACACATAGAGATCAATTTGCCAAATATTACAATCATGTAGTTCTAGTTTACATGCCAAAGT
 CTTCCCTTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT
 CATTTTGCAAGTAAAGAGCAACGGGACCCTTTCTAAAAACGTTGGTTGAAGGACCTAAATAC
 CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT
 TTCTCAGACACAACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT
 AATAATCTTTTGATGTTTTTAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA
 AGTCATTTAAACAAGCCACGGTGGGGCTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT
 GTCATTACTCCTGAATTATTACATTTTGGAGAATAAGAGGGCATTTTTATTTTATTAGTTACT
 AATTCAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC
 CAGATTGTCAGTGAAGCTGATGCCTAGGAACCTTTTAAAGGGATCCTTTCAAAGGATCACTT
 AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC
 AGTAATATATAAGTCACCTTACAGTGCTACTTCACACTTAAAAGTGCATGGTATTTTTTCATG
 GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA
 AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG
 CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT
 CATTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA
 AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG
 GTGGAATTTGTAATTAAATAATTATTAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
 LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
 IFSNFSIITTALLFRIVLKRRNLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
 SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
 YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
 FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
 YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDETF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,
 298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAC
GTATGGACCCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAACAGTGGAATGGAA
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCACAATGTATACATTCCTGCTAGG
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG
AAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT
TTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA
TGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCTGCTTTCTTTATTTCC
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAAGAGGCGTCTAAA
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA
CTAAAACCTTA

FIGURE 11

CGGACGCGCTGGGCGGACGCGTGGGCGGACGCGTGGGGCCGGCTTGGCTAGCGCGCGGCCGCC
GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCCTGCGGGGCAGAGGAGCAT
CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGGGGTCTATGGCCAAAGGAGAAGGCGCCGAG
AGCGGCTCCGCGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA
GGTGAAGAAAGAACCGAAAAAGAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCTTCAGATCTAC
CTATTG**ATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC
CTGGGATGCCATCACAGACCCCCCTGGTGGGCCCTCTGCATCAGCAAATCCCCCTGGACCTGCC
TGGGTGCGCTTATGCCCTGGATCATCTTCTCCACGCCCCCTGGCCGTCATTGCCTACTTCCTC
ATCTGGTTCGTGCCCCGACTTCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCCT
CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCCTTGTTTCCAGG
ACTTCAATAGCTCTACAGTAGCTTTCACAAAGTGCCAACCATACACATGGCACCCTTCACAC
AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTG
TGCTGTCA**T**CCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCAGCAGTCTG
AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCA**T**GAGCCACGGCCCATACATCAA**ACT**T
ATTACTGGCTTCTCTTACCTCCTTGGCTTT**C**ATGCTGGTGGAGGGGA**ACT**TTTGCTTTGTT
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCAGAA**T**CTACTCCTGGCCATCATGCTCT
CGGCC**ACT**TTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGAAGAAGACAGCT
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA
CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC
TACCCTGGTCCATGCTGCCTGATGTCA**TT**GACGACTTCCATCTGAAGCAGCCCC**ACT**TCCAT
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTT**CACCA**AGTTTGCCCTCTGGAGTGTC
ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC
CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG
CTGGGCCTGCTGCTCTTCAA**AA**TGTACCC**CA**TTGATGAGGAGAGGCGGCGGCAGAA**TA**AGAA
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG
AGCTGGCTAGCATCCTC**TAG**GGGCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG
AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA
CTGAAGACTCAAGGAGGTGGCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG
TGGCCTCCTGCCTCCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTACAGAGCC
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA
TTAATGTTATTAATTTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEPMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG
SCPTSHTARPIGTCEFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL
GTAIQGQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV
ILILGVREQREPYEAAQQSEPIAYFRGLRLVM SHGPYIKLITGFLFTSLAFMLVEGNFVLFCT
YTLGFRNEFQNLLLAIMLSATLTIPWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI
ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKAL
QALRDEASSSGCSETDSTELASIL

12/330

FIGURE 13

GGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACCTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATGT**
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA
GAGAACGTTATCATCAAATTAACAAGGCTGGCCTTGTAAGTACTGAGTTGTTTAGG
ACTTTCTATTGTGGCAAACCTCCAGAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTTCTACCAAATG
CAGCCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG
AGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT
ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTTCTGACTTACATTCGTGA
TTTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATTT**TGA**TGAAAGGAT
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAGTTGCTTA
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA
TAAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACTATG

Variable	Mean	SD	Min	Max
Age	34.5	10.2	22	55
Gender	0.45	0.50	0	1
Marital status	0.60	0.49	0	1
Education	12.5	1.5	9	16
Income	15.2	8.5	5	35
Health status	0.75	0.43	0	1
Employment	0.65	0.48	0	1
Stress level	3.2	1.8	1	5
Life satisfaction	4.1	1.2	2	5
Depression score	2.8	1.5	1	5
Resilience score	3.5	1.0	2	5
Optimism score	3.8	1.1	2	5
Self-efficacy score	3.6	1.0	2	5
Perceived social support	3.9	1.1	2	5
Life events score	2.5	1.5	1	5
Health-related quality of life	3.7	1.0	2	5
Overall well-being	3.3	1.1	2	5

Variable	Mean	SD	Min	Max
Age	34.5	10.2	22	55
Gender	0.45	0.50	0	1
Marital status	0.60	0.49	0	1
Education	12.5	1.5	9	16
Income	15.2	8.5	5	35
Health status	0.75	0.43	0	1
Employment	0.65	0.48	0	1
Stress level	3.2	1.8	1	5
Life satisfaction	4.1	1.2	2	5
Depression score	2.8	1.5	1	5
Resilience score	3.5	1.0	2	5
Optimism score	3.8	1.1	2	5
Self-efficacy score	3.6	1.0	2	5
Perceived social support	3.9	1.1	2	5
Life events score	2.5	1.5	1	5
Health-related quality of life	3.7	1.0	2	5
Overall well-being	3.3	1.1	2	5

FIGURE 16

CGGACGCTTGGGCNCGGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT
TACCAAGAATCTTCAACCCTTTCCCACAAAAGCTAATTGAGTACACGTTCCCTGTTGAGTACA
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA
GTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

FIGURE 17

CCCACGCGTCCGCCCCGCCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG
CCGGGGTGCGGAGCCGAC**ATG**CGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC
CTTCGCCTTGTA CT TGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG
AGGCTGGAGGCAGGTCGCTGTGGTTCCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG
GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCTTGCTCTTCTGCGGCGCCTA
CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT
TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC
TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCTGATAAAGT
GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTCTTATTGTTTT
TGAGACTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT
CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTTCATCTGTGT
GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG
TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT
AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA
CACAT**TGA**TCTGGATTTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA
TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG
CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT
TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC
ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAAT
ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC
AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT
CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRESEVLREYR
 KEHQAYVFLLEFCGAYLYKQGFAIPGSSFLNVLGALFGPWLGLLLCCVLTSVGATCCYLLSS
 IFGKQLVVSYPDPKVALLQRKVEENRNSLFFFLFLRLFPMTPNWFLNLSAPILNIPIVQFF
 FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKFSQKHLQ
 LNETSTANHIHSRKDT

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

FIGURE 19

CCGAGGCGGGAGGAGCCCCGAGGGGGCGCGAGCCCCGCATGAATCATTTGTAGTCAATCATTTT
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAAGTCCAGCTCAGAA
TAGGAAAATAACTTGGGATTTTATATTGGAAGACATGGATCTTGCTGCCAACGAGATCAGCA
TTTATGACAACTTTCAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT
TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA
GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTGCC
AACTGCACTGGCTGTGCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC
AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG
TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCCCTTATCCATGGAGGAGACCTCTGAA
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCACTCACCTGCCATTTCCAAAAGATG
CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG
ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA
GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCTG
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT
GGAACCGCTTTCTCAGAACTGTAGGAAATAGAAGTGTGCACAGGAACAGCTTCCAGAGCCGA
AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAAA

From

[illegible]

FIGURE 21

CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTCAAGTTGTGATC
AAGGGACACGTGGTTTCCGAACCTGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAAACTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAGAGAAGGCAATTGAAAATTTATC
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCGCAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

FIGURE 22

CCCACGCGTCCGCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG
 CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTG
 CCACTGGTGCGCACGCTGCTAGACCGTGCCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC
 CTCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA
 CACCCGAATGGCGCCACTTCATCGACAAACAGGTACAGCCAACC**ATG**TCCCAGTTCGAAATG
 GACACGTATGCTAAGAGCCACGACCTTATGTTCAGGTTTCTGGAATGCCTGCTATGACATGCT
 TATGAGCAGTGGGCAGCGGCCAGTGGGAGCGCGCCCAGAGTCGTCGGGGCTTCCAGGAGC
 TGGTGCTGGAACCTGCGCAGAGGCGGGCGCGCTGGAGGGGCTACGCTACACGGCAGTGTCTG
 AAGCAGCAGGCAACGCAGCACTCCATGGCCCTGCTGCAGTGGGGGGCGCTGTGGCGCCAGCT
 CGCCAGCCCATGTGGGGCCTGGGCGCTGAGGGACACTCCCATCCCCCGCTGGAACTGTCCA
 GCGCCGAGACATATTCACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCTCAC
 CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCTGACACCCACCGAGGAGGC
 CTCACCTGCTTGGCAGTGACCAAGAGGCCAAAGTGAGCACCACCCAGGTTGCTGCAGG
 AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAATG
 GATGAGCAGCGTGAGAAGCTGGTGTCTCGGCCGAGTGCCAGCTGGTGACGGTAGTGGCCGT
 GGTCCCAGGGCTGCTGGAGGTCACCACACAGAATGTATACTTCTACGATGGCAGCACTGAGC
 GCGTGGAACCCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCAGCTGCGTGAG
 GTCCACCTGCGGCGTTTCAACCTGCGCCGTTGACCACTGAGCTCTTCTTTATCGATCAGGC
 CAACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCCACTCATCTCTTAGCCAGA
 CTCCGAGACCCCAAGCCTGGCCCCATCCCACCCCATACCCAGGTACGGAACCAAGGTGTACTCG
 TGGCTCCTGCGCCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT
 GCTGCGTGCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACCTCGAGTACT
 TGATGCAACTCAACACCAATTGCGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTT
 CCCTGGGTCTGACGAGTACGTGTCCCCAACCCCTGGACCTCAGCAACCCAGCCGTCTTCCG
 GGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCAGCTCGTGAGGGAGAAGT
 ATGAAAGCTTTGAGGACCCAGCAGGGACCAATTGACAAGTTCCACTATGGCACCCACTACTCC
 AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGGAGCCCTTCACCTCCCTGCACGTCCA
 GCTGCAAGGTGGCCGCTTTGACTGCTCCGACCGGAGGAGCCCTCAATGTCTTCTATTACTGCACCTATGA
 GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTA
 TCAGCAACTTTGGGCAGACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACCTCGGCTCTCA
 GCTGAGGAAGCAGCCCATCGCCTTGACAGCCTGGACACTAACTCACCTAGCATCTTCCAGCA
 CCTGGACGAACCTCAAGGCATTCTTCGCAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
 CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTACGCTTACGCAAAAGACCCC
 ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGT
 GAGTGGACAAGCACTGGCAGTGGCCCCGATGGAAAGCTGCTATTACGCGGTGGCCACTGGG
 ATGGCAGCCTGCGGGTGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC
 CTTGATGTAGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCCG
 GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTGGCCTGGCAC
 CAAAGCCTGTGCAGGTCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT
 GAACCTTGACATGGCTGTGTCTGGATCTGAGGATGGAAGTGTGATCATAACACTGTACGCCG
 CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC
 TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTTGGGGCC
 CAGGTACCTACTCCTTGACCTGTATTAGTCAATGGGAAGTTGCGGGCTTCACTGCCCCCT
 GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTGTGTGCTGGGCACCGCCAGT
 GCGCCCTGCACATCCTCCAACATAACACACTGCTCCCGGCCGCGCCTCCCTTGCCCATGAAG
 GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCCTGGAGGA
 TGGCAAGCTCATCGTGGTGGTGGCGGGGCGAGCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGC
 GGAAGCTGTGGCGGTCCTCGCGCGCATCTCCAGGTGTCCTCGGAGAGACGGAATACAAC
 CCTACTGAGGCGCGCT**TGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCGCCGAGGCTG
 GCCCCGGAGGCCCCGCCAGAAGTGGCGGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGA
 GCGGGGCCACCCCTGCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCG
 GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGCCCGCCTGAGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL
 RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN
 HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELALELETP
 MEAAELDEQREKLVLSAECQLVTVVAVVPGLLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP
 LAQLREVHLRRLRNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV
 RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRITYNDL
 SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH
 YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKELEIP
 EFFYFPDFLENQNGFDLGCLQLTNEKVGDDVLPWASSPEDFIQQHRQALESEYVSAHLHEW
 IDLIIFYKQQRGPAAEEALNVFYYCTYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLEP
 HPTRLSAEEAAHRLARLDTNPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF
 SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGSLRVTALPRGKLL
 SQLSCHLDVVTCLALDTCGIYLISGSRDTCMVWRLLHQGGLSVGLAPKPVQVLYGHGAAS
 CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA
 WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA
 PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS
 GETEYNPTAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,
 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,
 873-879, 912-918, 954-960

FIGURE 24

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC
CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC
TGGACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC
GCACACTCCGTTACCACACTGGGTCATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAG
ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA
AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCCTGTCCTTCTTTT
TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC
TGGCTGCCCATCATGACCTCCATCCTGGGGGCTATGTCATCGCCAGCGGCTTCTTCAGCGT
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAG**TGA**CAGCTCCGGCCCTGATCCAGGACTGC
ACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT
AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC
GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA
AAAGATTTTATTAAAGATATTTTGTTAACTC

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLGF
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTIVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVLDKVTDLLLFFGKLLVVGGVGVLSEFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYSKSLKILGKKN
EAPPDNKKRKK

FIGURE 26

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT
 GGCT**ATG**TTTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC
 TTCTCTTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTTC
 CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAACTGC
 ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG
 ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG
 CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA
 TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT
 CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA
 GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT
 CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG
 CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC
 CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTGACGCG
 CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA
 CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC
 AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA
 GCGGCTCCAGGAGTTCCCTGTCAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC
 AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA
 TTTGGGATGAAGGACATGCGCGTGCAGACTTTCAGCATTCATTTTGGGTTCAAGCACAAAGTT
 TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT
 CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG
 TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC
 CTTTGACCAACCTCGTCATCTCCAGGGGCCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC
 TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA
 AGTCCTTTGTGTGTTTCGACAAAGAACCGGCGCTGCAAACCTGCTGCCCCCTGGTGATGGCTGCC
 CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC
 GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA
 TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT
 CTGGACGCACTTATTTCCCTCCTGTCC**TAG**GAATTTGATTCTTCCAGAATGACCTTCTTATT
 TATGTAACCTGGCTTTTCAATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT
 TTTTATTAAATAAAATGCTTATTTTAGGAAA

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAFL
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAIEDIFRDEEEDEEHSGNDS DGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGV LQRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR
LQEF LADMGLPLKQVKQKFQAMD ISLKENLREMI EESANKFGMKDMRVQTF SIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSNL DKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKS FVCSTKNRRCKLLPLVMAAP
LSMEHGT VTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNF DLSVIELKAEDRSKFL
DALISLLS

FIGURE 28

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGTGGCTATGNTCGTGTCCGATTTCCGCA
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT
GGTTCAGTTTCTGGGTGGCAAGAACTTGAAACTGCATTTCTTGAGCATAAAGAACAGTTTC
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA
CGATACCC

CACGAAACCCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTTCTAGTGAAC
CACGAAGGGACGATACACAGAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG
GCTGACTTTTGGCTATAGAAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAA
GTCTTCCCTTTATCGAGTCAAGAAACCCCCCTTCTTGAGCTATTTACAGCTTTTAAACAATT
GAGTAAAGTACGCTCCGGTACCA**ATGGT**GACAGCGCCCTGGGTCCCCTCTGGGCAGCGCTC
CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGAGCTCACCTTTGACAGAGCTGTGGC
CAGCGGCTGCCAACGGTGTCTGTGACTCTGAGGACCCCCCTGGATCCTGCCCATGTATCCTCAG
CCTCTTCTCTCCGGCCGCCCCACGCCCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC
CTGAAGGGTGACAAAGGGGACCCAGGCCCCAATGGGCTGCCAGGGTACATGGGCAGGGAGGG
TCCCCAAGGGGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCCG
GCGCCCCGTGCCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCCGCAAGACGGCCCTGCACAGC
GGCGAGGACTTCCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA
CATGGCGACCGGCCAGTTTGCTGCTCCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGC
ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTCTATC
CTGTACGCGCAGCCACGCGAGCGCAGCATCATGTCAGAGCCAGAGTGTGATGCTGGACCTGGC
CTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCAGAGAACGCCATCTACAGCA
ACGACTTCGACACCTACATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGAC**TG**AGGG
CCTCTGGGCCACCCTCCC GGCTGGAGAGCTCAGGTGCTGGTCCCCTCCCCTGCAGGGCTCAG
TTTGACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTTCTGGGGAGA
CCCTGCTTCTATCTTGGCTGCCATCATCCTCCAGCTATTTCTGCTCCTCTCTCTCTCTCT
TGGACCTATTTTAAGAAGCTTGCTAACCTAAATATTCTAGAACTTTCCAGCCTCTGTAGCCC
AGCACTTCTCAAACCTTGAAATGCATGCGAATCACCCGGGGTTCTGTGTTAAATGCAGATTCT
GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTTCTCATATGTTCTTGGGTGATGCTG
ATGGGGTCACTCTATGAACCACACTGGAGCAACCAGGTTCATTGACTTTCTCAATATTCTAG
TACTTCTGAACATTCTGGAATCTTCCACATTTCTAGAACTTCTCCCAACATTTTTTTTCT
TGAGACAGAGTCTTGCTCTGTTGCCCAGGCTAGAGTGCAGTGGTGCAATCTCAGTTCACTGC
AACCTCTGCCTCCCGGGTTCAAGCGATTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTAC
AGGCGCCTGCTACCATGCCTGGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATA
TTGGCCAGGCTGGTCTGAACCTCTGACTTCAGGTGACCCACCCGCTCTGGCCTCTCAAAT
GCTGGGATTACAGGTGTGAGCCACCGCTGCCTGGCCAATTCCAACATTTCTAAATCTCTCAT
CCCTCCAGGGCTCCCCGTGCTATGTTCTCTTTACCCCTTCCCCCTCTTCTCTTGCTCAGGCC
TGCACCACTGCAGCCACCGTTTCAATTTATTCATTCAATTAACACTGAGCACTCACTCTGTGCT
GGGTCCCGGGAAGGGTGAGGGGGTCAGACACAGGCCCCCTGCCCTCAGTGACTGGCCA
GTCCAGCCAGGCGGGGAGAGATGTGTACATAGTTTAAAGCAGACCCAGAGCTCATGGGG
GCCTGTGTTCTGGGTGTTTCAGGTGCTGCTGGTCTCCATTACCCACTGCTCCCCAAGGCTGG
TGGGACGGGGTCCCGGTGGCAGGGGCAGGTATCTCCTTCCCCTTCCCTCATCCACCTGCCAG
TGCTCATCGTTACAGCAAACCCAGGGGGCCTTGGCCAGGTCAAGGGTTCTGTGAGGAGAGG
ACCCAGGAGTGTGGGGGCAATTTGGGGGGTGAAGTGGCCCCCGAAGAATGGAACCCACACCCA
TAGCTCTCCCCACAGCTGATACGCGATCCTGCGAGAAGACTGCCCCCTCACTAGGGATCCC
CTTCTGCTCCTCTCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCACTCT
GAACTTCCGTTTCCCCAGGGCCTCCAGCTGCCCTCAGACACTGATGTCTGTCCCCAGGTGCT
CTCTGCCCTCATGCCCCCTCTCACCGGCCAGTGCCCCGACTCTCCAGGCTTTATCAAGGTG
CTAAGGCCCGGGTGGGCAGCTCCTCGTCTCAGAGCCCTCCTCCGGCCTGGTGCTGCCTTTAC
AAACCTTCAGAGAGAAGGGCCACGGAGCCAGGCTTTAGAGCCCTCAGCAGGTCTGGGG
AGCTAGAGCAAAGGAGGACCTCAGGCCATTCCGTTTCTTCTTCCAGGGTGGGGTGGCCTGGT
GTTCCCTTAGCCTTCCAACCCAGGTGGCCTGCCCTTCTCCCCAGAGGGAGGCGGCCTCCGC
CCATTGGTGCTCATGCAGACTCTGGGGCTGAGGTGCCCGGGGGGTGATCTCTGGTGCTCAC
AGCCGAGGGAGCCGTGGCTCCATGGCCAGATGACGGAACAGGGTCTGACCAAGTGCCAGGA
AGACCTGTGCTATAAACCAACCTGCCTGATCCTGCCCTGCCTGACCCGCCACGCCCTGCC
GTCCAGCATGATTAAAGAATGCTGTCTCCTCTTGGAAAAAAAAAAAAAAAAAA

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVFNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

C1q domain proteins.

amino acids 144-178, 78-111 and 84-117

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCCGACCCGCCAGGAAAGACTG
 AGGCCGCGGCCTGCCCCGCCGGCTCCCTGCGCCGCCGCCCTCCCGGGACAGAAG**ATG**TG
 CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG
 GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGG
 ACCACGGTGGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT
 CACCATGCTCGACGCAGGCAGCTTTGCCGGCTGCCGGGCCTGCAGCTCCTGGACCTGTAC
 AGAACAGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG
 GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCT
 CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC
 TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC
 CTGCCCCGCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT
 CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG
 AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACAGCTGGAG
 CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACAC
 CCGCATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG
 TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCGCTGCGG
 CTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCCTG
 GGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCGCCCCA
 AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAGCCACCACC
 ACCACAGCCACAGTGCCACCACGAGGCCCCGTGGTGCGGGAGCCACAGCCTTGTCTTCTAG
 CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCCAGCCCGCCCTCCA
 CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCACCCTCCACCTGCCTC
 AATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGCTTGTGCCCCGAAGGCTT
 CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA
 CGCCGAGGCCACCACGTCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC
 GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTACCTA
 TCGCAACCTATCGGGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTG
 AGTACACGGTCACCCAGCTGCGGCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGG
 CCCGGGCGGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGGCCATACACCCCCAGCCGTCCA
 CTCCAACCACGCCCCAGTCAACCAGCCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCCG
 CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG
 GGGCGGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCT
 GGAAGTGGAGGGAGTGAAGGTCCCCTTGAGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG
 AGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCAGGGCCTGGCCTC
 CAGTCACCCCTCCACGCAAAGCCCTACAT**CTAAG**CCAGAGAGAGACAGGGCAGCTGGGGCCG
 GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC
 CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA
 CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCCCTAACGTCCCCAGAAC
 CGAGTGCCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG
 GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCCACTCCAGGCGGA
 CCCTGGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG
 TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGCTTTA
 GGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCATTTATTCTG
 GGAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATG
 AAGGCCTTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 32

MCSRVP L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N
G I T M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R
R L E R L Y L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P
G I L D T A N V E A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G
N T R I A Q L R P E D L A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G
P W V R E S H V T L A S P E E T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S
S S L A P T W L S P T A P A T E A P S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E
G F T G L Y C E S Q M G Q G T R P S P T P V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L
T Y R N L S G P D K R L V T L R L P A S L A E Y T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A
V H S N H A P V T Q A R E G N L P L L I A P A L A A V L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G
P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E C E V P L M G F P G P G L Q S P L H A K P Y I

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
21

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT
CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**AGACTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAAGTGTATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTTCA
TTTTCAATCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCACAGCTGG
ATCTCAAAACCTTTGAATTCACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG
AAGAGTGTAACCTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAATGACTT
TGACACCATGCCTATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGA
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC
TTCTTAGGATTCAGAACTCTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC
AAAACCTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTCTTTTGGCTGATGGAATCA
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA
ATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTTGGCATAACATCAGTGGAACT
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC
TCAAATACTGTAATGAGAACTATAAAATGGAGCATGTACATTTAGAGTGTTTTACATTC
ACAGGATAAAATCTATTTGCTTTTGGCAAAATGGACATAGAAAACCTGACAATATCAAATG
CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATCCAATATTTAAATTTTGCC
AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACCTGCCTCACTTGAAAACCT
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC
CCTTGGAAACACTTGGATCTGAGTCAAAATCTATTACAACATAAAATGATGAAAATTTGCTCA
TGGCCAGAACTGTGGTCAATATGAATCTGTATACAATAAAATTTGCTGATTCTGTCTTCAG
GTGCTTGCCCCAAAAGTATTCAAATACTTGACCTAAATAATAACCAAATCCAAACTGTACCTA
AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAACTGAT
CTCCCTGGATGCAGTCATTTAGTAGACTTTAGTTCTGAACATTGAAATGAACTTCATTCT
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAA
ATCCATTCGGGTGATACCTGTGAATTAATAAAATTTTCAATTCAGCTTGAAACATATTCAGAGGTC
ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAAGTAG
GTTAAAGACGTTTATCTCCACGAATTATCTTGCAACACAGCTCTGTTGATTGTCAACATTG
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA
AGAACAACCTCAAGAGAAATGTCCGATTCCACGCATTTTATTCATACAGTGAACATGATTCTC
TGTGGGTGAAGAATGAATTGATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTGCT
CTTTATGAAAGCTACTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGAGTGGTGCCATT
ATGAATTTACTTTTGCCACCAATCTCTCCATGAAAATTTCTGATCATATAATTCTTATC
TTACTGGAACCCATTCCATTCTATTGCATTTCCACAGGTATCATAACTGAAAGCTCTCCT
GGAAAAAAAGCATACTTGAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAACTCTCTCTGATGAGAACAGATTGTCT
ATAAATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA
CAACCTTTATGATGGCAATTTGACAATATTTATTTAAAATAAAAAATGGTTATTTCCCTTCATA
TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG
GCTTATGGAAAAGGTGTTTATCCCAGGATTGTTTATAATCATGAAAATGTGGCCAGGTGC
AGTGGCTCACTCTTGTAATCCCAGCACTATGGGAGGCCAAGGTGGGTGACCCACGAGGTCAA
GAGATGGAGACCATCTGGCCACATGGTGAACCTGTCTCTACTAAAAATACAAAAATTA
GCTGGGCGTGATGGTGCACGCCTGTAGTCCCAGTACTTGGGAGGCTGAGGCAGGAGATCG
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAATGGAAAACATCC
TCATGGCCACAAAATAAGGTCTAATTCATAAATTATAGTACATTAATGTAATATAATATTA
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTGGTATGGAAAAACATATTAATAT
GTTATAAACTATTAGGTTGGTGCAAAATTAATTGTGGTTTTTGGCATTGAAATGGCATTGAA
ATAAAAGTGTAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTTGGA
TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTGTTTAGAATGA
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTkMDIENLTISNAQMPHMLFPNYPTKFQYLNfANNILTDELfKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSONLLQHKNDEncSWPETVVMNLSYNKLS
DSVFRCLPKSIQILDlnnnQIQTVPKETIHLMAALRELNIafNFLTDLPGCSHFsrLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFCCCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQlKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHNLfHENSdHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRCGLFWANLRAAINVNVLATREMYELQTfTELNEESRGSTISLM
RTDCL

FIGURE 35

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLET LKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPMPVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNI GKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNP DGYEKAYEGG
SELGGWSLGRWTHDGI DINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTL SKTNMARIREIMEKFGKQPVSLPARR
LKLGRGRKRRQRG

FIGURE 37

CTAAGAGGACAAG**ATG**AGGCCCGGCCTCTCATTTCTCCTAGCCCTTCTGTTCTTCCTTGGCCAAGCTGCAGGGG
 ATTTGGGGGATGTGGGACCTCCAATTTCCAGCCCCGGCTTCAGCTCTTCCAGGTGTTGACTCCAGCTCCAGC
 TTCAGCTCCAGCTCCAGGTGGGCTCCAGCTCCAGCCGAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT
 TTCCAATTTACCCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC
 CCGTGGACAGAGTGGAACGCTTGAATTCACAGCTCATGTTCTTTCTCAGAAGTTTGAGAAAGAACTTTCTAAA
 GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTAAACCTAACTGTCCGAATTGACATCAT
 GGAGAAGGATACCATTTCTTACACTGAAGTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC
 TGGTCATACAGCTGAAGGAGAGTTTTGGTGGGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT
 ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTGCGCCGAGAAATCGTGGC
 TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTGTCACCCCTCTCCCACTC
 CAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTTT
 TCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCCAGCATCCAAACAAAGGACTGTATTGGGTGGCGCC
 ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTGTCTATTGTATATAA
 ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC
 ATGTACAACACCGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACCTCTCCCTAA
 TGCTGCCTATAATAACCGCTTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAAATG
 GATTGTGGGTATTTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT
 CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCCTTCATGGTATGTGGGGTTCT
 GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAGAGG
 GCAAACCTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGACAGCATTAACCTATAACCCCTTTTGACCAGAAA
 CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAG**TAA**GCTGTTTA
 GGAGTTAGGGTGAAAGAGAAAAATGTTTGGTGAATAATAGTCTTCTCCACTTACTTAGATATCTGCAGGGGTGT
 CTAAAAGTGTGTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT
 TGATTTGGTGAGTTCTCTTGGGAATCATCTGCCTCTTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGA
 TTGTGAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA
 GGAATTAAGGAACCTTAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC
 CTCATCCATGTAGCACCCTAATTCTTCCATGCCTGGAAGAAACCTGGGGACTTAGTTAGGTAGATTAATATCT
 GGAGCTCCTCGAGGGACCAATCTCCAACCTTTTTTTTCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG
 CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTT
 ATGCATTAAATTTGATAGTGGCAAATAAATCCCAGAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTCTCTC
 ATTTGTCCACCTTACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCCTTCCAAAGGCAGCTCAGAAGATTAG
 AACCAGACTTACTAACCAATTCCACCCCCACCAACCCCTTCTACTGCCCTACTTTAAAAAATTAATAGTTTT
 CTATGGAAGTATCTAAGATTAGAAAAATTAATTTTCTTTAATTTTATTATGGACTTTTATTTTACATGACTCTA
 AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA
 TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTGCCTTTGTTTAAGCCTGGAACCTGTAAGAAAAAT
 GAAAAATTTAATTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT
 TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTTT
 TCCTTTGATGTTCAAGTCCTAGTCTATAGGATTGGCAGTTTAAATGCTTTACTCCCCCTTTTAAAAATAAATGAT
 TAAATGTGCTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 38

MRPGLSFLLALLFFLGQAAGDLGDVGPPIPSPGFSSFPGVDSSSSFSSSSSRSGSSSSSRSLGS
GGSVSQLFSNFTGSVDDRGTCCQCSVSLPDTTFPVDRVERLEFTAHVLSQKFEKELSKVREYV
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTKLKECEASKDQNTPVVHPPPTPGSCGH
GGVVNISKPSVVQLNWRGFSYLYGAWGRDYSPQHPPNKGLYWVAPLNTDGRLLLEYRLYNTLD
DLLLYINARELRITYGQSGTAVYNNMYVNMVNTGNIARVNLTNTIAVTQTLPNAAAYNNR
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDTTLQVLNTWYTKQYKPSASNAF
MVCGLYATRMTNTRTEEIFYYYDTNTGKEGKLDIVMHKMQEKVQSINYNPFDQKLYVYNDG
YLLNYDLSVLQKPQ

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCCTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTACAACACCGGGNATATTGCCAGAGTTAACCTGACC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189

FIGURE 40

TCTCGCAGATAGTAAATAATCTCGGAAAGGCGAGAAAGAAGCTGTCTCCATCTTGTCTGTAT
 CCGCTGCTCTTGTGACGTTGTGGAG**ATG**GGGAGCGTCCTGGGGCTGTGCTCCATGGCGAGCT
 GGATACCATGTTTTGTGTGGAAGTGCCCCGTGTTTGCTATGCCGATGCTGTCTAGTGGAAC
 AACTCCACTGTAACTAGATTGATCTATGCACTTTTCTTGCTTGTGGAGTATGTGTAGCTTG
 TGTAATGTTGATACCAGGAATGGAAGAACAACGAATAAGATTCCTGGATTTTGTGAGAATG
 AGAAAGGTGTTGTCCCTTGTAAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT
 GGTGGCTATGTTCTATCTTCTTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGA
 TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA
 TTATTGGGGCATTCTTCATTCCAGAAGGAACCTTTTACAACGTGTGGTTTTATGTAGGCATG
 GCAGGTGCCTTTTGTTCATCCTCATACAACCTAGTCTTACTTATTGATTTTGCACATTCATG
 GAATGAATCGTGGGTGAAAAAATGGAAGAAGGGAACTCGAGATGTTGGTATGCAGCCTTGT
 TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC
 TACACTCATCCAGCCAGTTGTTTCAGAAAAACAAGGCTTCATCAGTGTCAACATGCTCCTCTG
 CGTTGGTGCTTCTGTAATGTCTATACTGCCAAAAATCCAAGAATCACAAACCAAGATCTGGTT
 TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT
 GAACCAGAAACAAATTGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC
 TGTCCCAAAGGAAGGGCAGTCACTCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC
 TCTTTTGTGTGTGATTTTATTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA
 CTGACTCTAACAAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC
 ACTGGAGGATGGGGACGATGTTTACCAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA
 GTTATTCCTTCTTTCATTCATGCTTTTCTGGCTTCACTTTATATCATGATGACCCCTTACC
 AACTGGTCCAGGTATGAACCCTCTCGTGAGATGAAAAGTCAGTGGACAGCTGTCTGGGTGAA
 AATCTCTTCCAGTTGGATTGGCATCGTGCTGTATGTTTGGACACTCGTGGCACCAGTTGTTT
 TTACAAATCGTGATTTTGACT**GA**GTGAGACTTCTAGCATGAAAGTCCCACTTTGATTATTGC
 TTATTTGAAAACAGTATTCCCAACTTTTGTAAAGTTGTGTATGTTTTTGTCTCCCATGTAAC
 TTCTCCAGTGTCTGGCATGAATTAGATTTTACTGCTTGTCATTTTGTTATTTTCTTACCAA
 GTGCATTGATATGTGAAGTAGAATGAATTGCAGAGGAAAGTTTTATGAATATGGTGATGAGT
 TAGTAAAAGTGGCCATTATTGGGCTTATTCTCTGCTATAGTTGTGAAATGAAGAGTAAAA
 ACAATTTTGTGACTATTTTAAAAATTATATTAGACCTTAAGCTGTTTTAGCAAGCATTAAA
 GCAAATGTATGGCTGCCTTTTGAATATTTGATGTGTTGCCTGGCAGGATACTGCAAAGAAC
 ATGTTTTATTTTAAAAATTTATAAACAAGTCACTTAAATGCCAGTTGTCTGAAAAATCTTATA
 AGGTTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGG
 TTATGGATGGAGGTGTCGGTACTAAATTTGAATAACGAGTAAATAATCTTACTTGGGTAGAGA
 TGGCCTTTGCCAACAAAGTGAAGTGTGTTTGGTTGTTTTAAACTCATGAAGTATGGGTTCAGT
 GGAAATGTTTGGAACTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGA
 AGGAAGTGTGTTGAAAGTCACTTTGAAAGTTAGTTTTGGGCCAGCACGGTAGCTCACCTT
 GGTAATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCAGGAATTCAGACCA
 GCTTGGCACATGGTGAACCTGTTCTATAAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC
 CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCAGAGCCAAAGGTTGCAGTGAGCAAGTCA
 CGTCACTGCACTCTAGCTGGCACAGAGTAAGCCAAAAAATATATATATATTGAAATCAAGG
 AGGCAAAATTTTGGACAGGGAAGGAAGTAACTGCAAAACCACTAGGCTTTAGTAGGTACTTAT
 ATAAAATCTAGTCCAGTTCTCTCATTTAAAAAATGAAGACACTGAAATACAGACTTAAATA
 GCTCAGATAGCTAATTAGGAAATTTCAAGTTGGCCAATAATAGCATTCTCTCTGACATTTAA
 AAATAATTTCTATTCAAAATACATGCATATTGATTTACACCTCATACTGTGATAATTAATGT
 GATGTGGATTGCTGGTGTCCAGCATGACCCATAAACAGGTCAGAAGAATGATGGAATGTTTT
 AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAAGATGTTTAAAAATGCTTTTGTAT
 TTAAGTCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT
 ATGAGAGTTTAGTTATTTGTATGTGTCACTAGTGCTAATGAAGCTTTTAAAAATCTACAATT
 TCTTCTTTAAAAATATTTATTAATGTGAATGGAATATAACAATTCAGCTTAATTCCCCAACCC
 TTATTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAA
 ATGAATTCAGAGAAAAA

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSNKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTSAMTNEPETNCNPSSLIIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVITYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVWKISSSWIGI
VLYVWTLVAPLVLNTRDFD

FIGURE 42

CGCAGAAAGAGCTGCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGTCTATGCCGATGCTGTCCTAGTGGAACAANTCCACTGTAAGTAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACAACCTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
GGGC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2

FIGURE 44

AAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGTGAACGTTNTGGAGATGGGGAGC
GTCCTTGGGGTTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCCCCGTGTT
TGCTATGCCGATGCTGTCCTAGTGGAAACAACCTCCACTGTAACTAGATTGATCTATGCACTT
TTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAAC
GAATAAGATTCCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATTTTGGTTG
GCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTA
CTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTT
CTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
21

FIGURE 45

GCTGTCCTTAGTGGAACAANTCCA¹ACTTGTA²ACTTGGATTGATCTATGCACTTTTTCTTG
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA
AGATTCCTGGATTTTTGTGAGAATGAGAAAGGTGTTGTCCCCTTGTAACATTTTTGGTTGGC
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTACT
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACTTTT
ACAAC³TGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGT⁴TTTCATCCTCATACAAC⁵TAGT
CTTACTTATTGATTTTGCACATT⁶CATGGAATGAATCGTGGGTGAAAAAATGGAAGAAGGGA
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC
GTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

FIGURE 46

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCCTGCGATTGAGCTGCGGGTTCGCGCCGGCGCCGGCCTCTCCAAT
GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTCGGCAAAGGCAGTCGAGTGTTCGACACCGGGGCGAG
TCCTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCCGGCCGGGGCTGTGCG
ACTCCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAAGCGGAAAAGAGGCAGATTAC
GTCGTTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC
CCCTGGTTTTGTGTGTACGCACACACAGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCC
TGGGCGAATCCACATCTGTTTCACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTCGAATCTGCGAGTG
AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAGAACACCAGAT
CAGCAAAAAAAGAAGATGGGGCCCCCGAGCCTCGTGTCTGTCTGTCTGCTGTCGCAACTGTGTTCTCCCTGCTGGG
TGGAGCTCGGCCTTCCTGTGCGACCAACCGCTGAAAGGCGAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCCA
ACATCATCCTGGTGTGACGGACGACCAGGATGTGGAGCTGGGTTCATGCAGGTGATGAACAAGACCCGGCGC
ATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCCTCACGCTCCTC
CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC
AGGCACAGCAGGAGAGCCGCACTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG
TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTGGGACTCCTTAAAACTCCCG
CTTTTATAACTACACGCTGTGTGCGAACGGGTGAAAGAGAAGCAGCGCTCCGACTACTCCAAGGATTACCTCA
CAGACCTCATACCAATGACAGCGTGAGCTTCTTCGCGACGTCCAAGAAGATGTACCCGACAGGCCAGTCCCTC
ATGGTCATCAGCCATGCAGCCCCCAGGCCCTGAGGATTCAGCCCCACAATATTACGCCTCTTCCCAAACGC
ATCTCAGCACATCAGCCGAGCTACAACCTACGCGCCCAACCCGACAAACACTGGATCATGCGCTACACGGGGC
CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGCCCTCATGTGGTGGAC
GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACGGC
CGACACCGTTACCACATCGGCCAGTTTGGCCTGGTGAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG
TCCCGTTTCTACGTGAGGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG
GCCCCACCATCCTGGACATTGACGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCT
GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG
AGAGAGGCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAACTTTCTGCCCCAGTAC
CAGCGTGTGAAGGACCTGTGTGACGCTGCTGAGTACAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG
TGTGGAGGACGCCACGGGGAAAGCTGAAGCTGAAGCTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC
TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTC
AGCCTGGCCGGACGCCGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCGAGTCGCTCCATCCG
CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA
CCAAGCGGCACTGGCCAGGGGGCCCCGTGAGGACCAAGATGACAAGGATGGTGGGACTTCAGTGGCACTGGAGGC
CTTCCCGACTACTCAGCCGCCAACCCCATTAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA
GTGTGACCTGGACCTGTACAAGTCCCTGACGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAA
CCCTGCAGAAACAAATTAAGAACCTGAGGGAAGTCCGAGGTCACTGAAGAAAAAGCGGCCAGAAGAATGTGAC
TGTACAAAATCAGCTACCACACCCAGCACAAAGGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTTCAG
GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGC
TCAAGCGCTGCAGAACAACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG
CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCTGACACAGCGCCAACAATAACACGTACTGGTGCAT
GAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCA
ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCTCAACCAGCTACACGTACAG
CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAACCCCGGACTCGAAACATGGACCTGGATGGAGG
AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAATCACTGG
GACAACCTGTGGGAAGGCTGGGAAGGTTAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA
CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGCTGTGCTATTGGCCAGGAGGCCGTGAGAAAGC
AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC
ATTTTTTGCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTTCGTATCAAAAAGTC
ACCACTAACCTCCCCCAGAAGCTCACAAGGAAAAACGGAGAGAGCGAGCGAGAGAGATTTCTTGGAAATTTTC
TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCCTGTTCTAAATCCTCTATTCTT
TTGGTTTTGTACAAAAGGAAGTAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG
TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCTGGTTGCCTCTGAAGAAA
CTGCCTTCATTGTATATATGACTATTTACATGTAATCAACATGGGAACCTTTAGGGGAACCTAATAAGAAAT
CCCAATTTTCAGGAGTGGTGGTGTCAATAACGCTCTGTGGCCAGTGTAAAAGAAAAA

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHE
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV AIEVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQEKDKVWLLREQKRKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMRTINETHNFLCFEATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA
GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCC
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCTTGCTGCTGCTGCTGCTGCTGGCCACC
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGAAA
CCGAGTCCGCCGGGCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA
CGCTCGCTGAGGCTGCTGTGCGCCGGTGCCTGTGGACAGCAGCTGCCCCTGCCCTCCCATCTG
TCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGAAGTCCCAGTGAGCCCCAGAAATG
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTTCATCGGTGCCTTAGTCCAAGAAAAT
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

199

FIGURE 49

MLGLLGSTALVGWITGAAVAVLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRRAQWPWFR
RRGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPHRRHPRHAR

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTGGGGATCCAGAGCC**ATG**TCGGACCTGCTA
CTACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG
GTACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTCACCCCCATCCGCAACGTCA
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTTCTGTCCATCTGGCTGGCTACCCGCCG
TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG
AGATCTACCAGGAAGACCAGATCCATTTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTAT
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT
GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA
GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAG**TAA**CCC
ATGGCCTGCACCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT
CCAGCCCTCTTCCTCCTTCTCTGGGGGAGGAGGGGTTCTGAGGGACCTGACTTCCCCTGC
TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTCAGACTC
ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAA
AAAAA

FIGURE 51

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRAVGSILSEGEESPPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATTRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAAATLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

FIGURE 52

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCT
GCCCCGCGCCCAGTC**ATG**ACCCTGCGCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA
CCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGA
GACACGCTTCACATACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCCTTCTCACTTGGCCTAT
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT
GATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTTCCTCTGGTAG
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAAT
AGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAA**TA**
ATAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA

FIGURE 53

MTLRPSLLPLHLLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSKKK

MTLRPSLLPLHLLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSKKK

FIGURE 54

CCCGGGAACGTGTTCTGGCTGCCGCACCCGAACAGCCTGTCTGGTGCCCCGGCTCCCTGC
CCCGCGCCCAGTCATGACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGC
TGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGACC
CTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGAGA
CACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCTGA
CCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGT
CTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTATGG
AAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCTGA
TTGCACTAATCCGAGCCAACCTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAGGG
ATGGCCATGGTGCCACCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAATAGA
CCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATAATA
AATAATAAATTTTAAAAAACTTA

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAAATAATAAATTTTAAAAAACTTAAAA

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG
TGAGGCGGGCCGGCGCGGCGCGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG
ACCTGAAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG
GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT
TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT
GTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA
GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG
TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG
CTAAAGAAAAAGACATAGTATAACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT
TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAGTGAACACATCTGAT
TTCCACAGCACAACAGCCCTGCATGGGTTTGTGTTGTTTTTTTACTGCTCACTCCCAACCTT
TTGTAATGCCATTTTCTAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT
AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACCT
ATAAAATGTTAGAGGAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT
TGTAATAATAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA
TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT
CTCTTCTGTGTCAAATGTAAATGAAATATAAACATTTTTTTAGTTTTTAAATATTCCGTGG
TCAAATTTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT
GTAATGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT
GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCCTCTCATT
CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC
CACATCCACCACTG

57/330

FIGURE 57

MSGFLEGLRCSECIDWGEKRTIASIAAGVLFFTGWWIIIDAAVIYPTMKDFNHSYHACGVI
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

FIGURE 63

CGACGCCGGCGTG**ATGT**GGCTTCCGCTGGTGCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCC
 TCTGCAAAGTTTACTTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC
 AAACGGCCCCCAGCGCCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC
 TTTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG
 GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCTGGTGCTGGAACAACAT
 ACCAAGGCAGGGGGGCTGCTGTCATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT
 CCATTACATTGGGCGTATGGAAGAGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG
 AAGGGCAGCTGGACTGGGCTCCCCTGTCCTCTCCTTTTTGACATCATGGTACTGGAAGGGCCC
 AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTCAGGGCCTCAAGGA
 GAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA
 GTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCGTGGTTCAGCTCCTCGAC
 AGGTGTGGGCTGCTGACTCGTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGA
 GTCCTGTCAGCAGCTGGGGGCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA
 CTTACGGTGTGCACCCCCAACCCAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC
 ATGAAAGGAGGCTTTTTATCCCCGAGGGGGTTCCAGTGAAATTGCCTTCCACACCATCCCTGT
 GATTTCAGCGGGCTGGGGGCGCTGTCCTCACAAGGCCACTGTGCAGAGTGTGTTGCTGGACT
 CAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC
 CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC
 CCGCTGCCCTGCCAGGTGTGAAGCAGCAACTGGGGACGGTGCGGCCCGGCTTAGGCATGACCT
 CTGTTTTTCATCTGCCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT
 GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCTCCATGCCCAGGGGAAGA
 GGCTGCGGAACACATCCCTCTTCTCTTCTTCCGCTTTCCCATCAGCCAAAGATCCGACCTGGG
 AGGACCGATTCCCAGGCCGGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTTT
 GAGGAGTGGCAGGCGGAGCTGAAGGGAAAGCGGGCAGTGACTATGAGACCTTCAAAAACCTC
 CTTTGTGGAAGCCTCTATGTGAGTGGTCTGAAACTGTTCCACAGCTGGAGGGGAAGGTGG
 AGAGTGTGACTGCAGGATCCCCACTCACCAACCAGTTCTATCTGGCTGCTCCCCGAGGTGCC
 TGCTACGGGGCTGACCATGACCTGGGCCGCTGCACCCTTGTGTGATGGCCTCCTTGAGGGC
 CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTACCTGTGGACTGGTCG
 GGGCCCTGCAAGGTGCCCTGCTGTGACGAGCGCCATCCTGAAGCGGAACCTGTACTCAGAC
 CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAT**TAGT**TCCATCAGGGAGG
 AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
 CATTAGTTTCTTGACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
 TTTAAATCACAATTTCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
 TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATAGCTTGATG
 TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCTCCTAACTCAGTGATCAAAGCGA
 ATATTCCATCTGTGGATAGAACCCTGGCAGTGTGTGTCAGCTCAACCTGGTGGGTTCAGTTC
 TGTCCTGAGGCTTCTGCTCTCATTCATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG
 GAAAAGGGAGACTAATGAGGCTTAACCTCAAAACCTGGGCGTGGTTTTGGTTGCCATTCCATA
 GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTCAGTGGCTCTTCAGGGGACAGGAAAT
 GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA
 GGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCGGGTCTGTCTTCTTATCA
 TCGGGGTGGCAGCTGGTTCTCAATGTGCCAGCAGGGAAGTACCTGAGCCTCAATCAAGC
 CTTATCCACCAAATACACAGGGAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCA
 TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCAGCCAAGGG
 CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGAAGTCACATCAGAAAAGGGA
 AAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTTCAGTGAATTAGCACATGTGAGGG
 TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTG
 GTATCAGACATACGAAAGGTCTCTTTGTAGTTTCGTGTTAATGTAACATTAATAAATTTATTG
 ATTCCATTGCTTTAAAAA

FIGURE 64

MWLPLVLLLAVLLLAVLCKVYLGLFSGSSPNPFSEDKRPPAPLVTDKARKKVLKQAFSAN
QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLFDTGIHYIG
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ
EEAIIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ
LGASSELQAVLSYIFPTYGVTPNHSAFMSHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA
GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP
GVKQQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTDMDQAMERYVSMFREEAAEH
IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAELKGKRGSDYETFKNSFVEA
SMSVVLKLFPPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI
PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

FIGURE 65

[illegible]

FIGURE 66

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLD
SESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG
EPCHFPPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMQEAMM
YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMF EK
LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

67/330

FIGURE 67

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCCATGGCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCAGCACGTCAACCATCAGGGACTACGGTGTGTCCTGG
TACCAGCAGCGGGCAGGCAGTGGCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCAACAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCCTAGGGGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTCT
GCCCCTGACCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAATGGG
TTAATAATATTCAACATGTCAACAAC

FIGURE 69

GCCGCCCCGCCCCGAGACCGGGCCCCGGGGGCGCGGGGCGCGGGGATGCGGGCGCCCGGGGCGG
 CGATGACCGCGGAGCGCACGCCGCGGGCCCCGGCCCTGACCCCCGCCGCCCGCCCGCTGAGCCC
 CCCGCCGAGGTCCGGACAGGCCGAGATGACGCGGAGCCCCCTGTTGCTGCTCCTGCTGCCGC
 CGCTGCTGCTGGGGGCCTTCCCACCGGCCGCCGCCGCCGAGGCCCCCCAAAGATGGCGGAC
 AAGGTGGTCCCACGGCAGGTGGCCCCGCTGGGCCGCACTGTGCGGCTGCAGTGCCCAAGTGA
 GGGGGACCCGCCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA
 GCCGCTTCCGCGTGTGCGCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC
 GTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCCTCGTCGT
 GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCCGACAGCTCCTCTGGGGGTCAAG
 AGGACCCCCGCCAGCCAGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGC
 CGGGTGATCGCACGGCCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCGTGGCCAGCGGGCACCC
 TCGGCCCGACATCACGTGGATGAAGGACGACACAGGCCTTGACGCGCCAGAGGCCGCTGAGC
 CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGGAGGACAGCGGCAATAC
 ACCTGCCGCGTGTGCAACCGCGCGGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA
 GCGGACCCGTTCCAAGCCGCTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG
 GGGGGACCACGTCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG
 AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACCTCCACCATCGATGTGGGCGGCCAGAA
 GTTTGTGGTGTGCCCCACGGGTGACGTGTGGTGCGGCCCGACGGCTCCTACCTCAATAAGC
 TGCTCATCACCCGTGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGCGGCCAACACC
 ATGGGCTACAGCTTCCGCGAGCGCCTTCCTCACCGTGTGCGCAGACCCAAAACCGCCAGGGCC
 ACCTGTGGCCTCCTCGTCTCGGCCACTAGCCTGCCGTGGCCCGTGGTTCATCGGCATCCAG
 CCGGCGCTGTCTTCATCCTGGGCACCCCTGCTCCTGTGGCTTTGCCAGGCCCAGAAGAAGCCG
 TGCACCCCCGCGCTGCCCTCCCCTGCCCTGGGCACCGCCCGCGGGGACGCCCGCAGCCG
 CAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTGTGGGGCTGT
 GTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCAGGCCCAGTTGCTGGC
 CCTAAGTTGTACCCCAAACCTCTACACAGACATCCACACACACACACACACACTCTCACAC
 AACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGCTAGACGGCACCGT
 ATCTGCAGTGGGCACGGGGGGGCCGGCCAGACAGGACAGCTGGGAGGATGGAGGACGGAGCT
 GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCAGGAGTCTGTGTG
 TGAGGCATAGCCCCCTGGACACACACACACAGACACACACACTACCTGGATGCATGTATGCAC
 ACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGCACACGCACATGCACAGATATG
 CCGCCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGAGACATGCCAGAACA
 TACAAGGACATGCTGCCTGAACATACACACGACACCCATGCGCAGATGTGCTGGGACA
 CACACACACACACGGATATGCTGTCTGGACGCACACACGTGCAGATATGGTATCCGGACACA
 CACGTGCACAGATATGCTGCCTGGACACACAGATAATGCTGCCTTGACACACACATGCACGG
 ATATTGCCTGGACACACACACACACACACGCGTGCACAGATATGCTGTCTGGACACGCACAC
 ACATGCAGATATGCTGCCTGGACACACACTTCCAGACACACGTGCACAGGCGCAGATATGCT
 GCCTGGACACACGCAGATATGCTGTCTAGTCACACACACACGCAGACATGCTGTCCGGACAC
 ACACACGCATGCACAGATATGCTGTCCGGACACACACACGCACGCAGATATGCTGCCTGGAC
 ACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATATTGCCTGGACACACACA
 TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACG
 CACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACACAGGT
 GCAGATATGCTGCCTGGACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGA
 TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA
 TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA
 CACACATGCACACACAGGTGCAGATATGCTGCCTGGACACACGCAGACTGACGTGCTTTTGG
 GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT
 CCCTGCTCCACCGTCACTCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCTC
 CATCCCCGCTCTGTCCCCCTGGCCTTGGCGGCTATTTTTGCCACCTGCCTTGGGTGCCCAGG
 AGTCCCCCTACTGCTGTGGGCTGGGGTTGGGGGCGACAGCCCCAAGCCTGAGAGGCTGGAG
 CCCATGGCTAGTGGCTCATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCCTTGGA
 TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTGTCAGGGAC
 TGTGGTCTCTCCTGGGGCCCCGGGACCCGCTGGTCTTTTCCAGCCATGCTGATGACCACACCCC
 GTCCAGGCCAGACACCACCCCCACCCCACTGTCGTGGTGGCCCCAGATCTCTGTAATTTTA
 TGTAGAGTTTGAGCTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

70/330

FIGURE 70

MTPSPLLLLLLPLLLLGAFPPAAAARGPPKMADKVVPQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDDISPGK
ESLGPDSSSSGGQEDPASQQWARPRFTQPSKMRRRVIA RPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTVPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTSHTHSHVEGKV
HQHIHYQC

FIGURE 71

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAACTTCCCAGGGGACCGCATTCAGAGTC
 AGTGA CTCTGTGAAGCACCCACATCTACCTCTTGCCACGTTCCCACGGGCTTGGGGGAAAGATGTTGGGGACCA
 AGGCCTGGGTGTTCTCCTTCTGCTGGAAGTCACATCTGTGTGGGGAGACAGACGATGCTACCCAGTCA
 GTAAGAAGAGTCCAGCCTGGGAAGAAGAACCCAGCATCTTTGCCAAGCCTGCCGACACCCCTGGAGAGCCCTGG
 TGAGTGGACAACATGGTTCAACATCCGACTACCCAGGCGGGAAGGGCGACTATGAGCGGCTGGACGCCATTGCT
 TCTACTATGGGGACCGTGTATGTGCCCCGTCCCCTGCGGCTAGAGGCTCGGACCACAGTGGACACCTGCGGGC
 AGCAGTGGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTCTGGTGCCTCAACAGGGAGCAGCGGCTGGCCA
 GAAGTCTCTAATTACACCGTACGCTTCCCTCTGCCACACAGGATCCCTGCGCCGAGACACAGAGCGCATCTGGA
 GCCATGGTCTCCCTGGAGCAAGTGCTCAGCTGCCTGTGGTCAAGTGGGCTCCAGACTCGCACACGCATTTGC
 TTGGCAGAGATGGTGTGCTGTGAGTGGGCGAGCGAAGAGGGTCAAGTGCATGGGCCAGGACTGTACAGC
 CTGTGACCTGACCTGCCAATGGGCGGCTGTAAGTGTGACTGTGATGCCGTGCATGTGCCAGGACTTCATGCTTC
 ATGGGGCTGTCTCCCTTCCCAGGAGTGGCCAGCCTCAGGGGCTGCTATCTACCTCCTGACCAAGACGCGGAAG
 CTGCTGACCCAGACAGACAGTGTATGGGAGATTCCGAATCCCTGGCTTGTGCCCCTGATGGCAAAAGCATCCTGAA
 GATCACAAAGGTCAAGTTTGGCCCATTTGCTTACTCAAACTGCCAAGACTAGCCTGAAGGCGACCCACCATCAAG
 CAGAGTTTGTGAGGGCAGAGACTCCATACATGGTGTATGAACCCCTGAGACAAAAGCAGGAGAGTGGGCGAGAGC
 GTGTCTGTGCTGTGAAGGCCACAGGGAAGCCAGGGCAGACAAGTATTTTGGTATCATAATGACACATTTGCT
 GGATCCTTCCCTCTACAAGCATGAGAGCAAGTGGTGTGAGGAACTGCAGCAGCAGCCAGGCTGGGGAGTACT
 TTTGCAAGGCCAGAGTGTGCTGGGGCTGTGAAGTCCAAGGTTGCCAGCTGATTGTACAGCATCTGATGAG
 ACTCCTTGCAACCCAGTTCTGTAGAGCTATCTTATCCGGCTGCCCATGATTGCTTTCAAGATGCCACCAACTC
 CTTCTACTATGACCTGGGAGCGCTTGGCTTTAAGACTTGTGAGGGCAGCAGGATAATGGGATGGGCTGCTG
 ATGCTGTGCAAGTGTGTGGCATCTCCAAGACAGAGGAAAGGGAGATCCAGTGCAGTGGCTACACGTACCC
 ACCAAGGTGGCCAAAGGAGTGCAGCTGCCAGCGGTGTACGAAACTCGGAGCATCGTGGGGGCGGTGTCACTGC
 TGCTGACAATGGGAGGCCATGCGCTTTGGCCATGTGTACATGGGGAACAGCCGTGTAAAGCATGAGTGGCTACA
 AGGGCACTTTTCAACCTCCATGTCCCCCAGGACACTGAGAGGCTGGTGTCAATTTGTGGACAGGCTGCAGAA
 TTTGTCAACACCACCAAGTGTACCTTTCAACAGAGGGGAGTGGCGTGTCCATGAAATCAAGATGCTTCCG
 TCGGAAAGAGCCCATCACTTTGGAAGCCATGGAGACCAACATCATCCCCCTGGGGGAAGTGGTGAAGACC
 CCATGGCTGAAGTGGAGATTCCATCCAGGAGTTTCTACAGGCAGAAATGGGAGGCCCTACATAGGAAAAGTGAAG
 GCCAGTGTGACCTTCTGGATCCCCGGAATATTTCCACAGCCACAGCTGCCAGACTGACCTGAAGTTCATCAA
 TGACGAAGGAGACACTTTCCCCCTTGGACGTATGGCATGTTCTCTGTGGACTTCAGAGATGAGGTCACCTGAG
 AGCCACTTAATGCTGGCAAAGTGAAGGTCCACCTTGACTCGACCCAGGTCAAGATGCCAGAGCACATATCCACA
 GTGAAACTCTGGTCACTCAATCCAGACACAGGGCTGTGGGAGGAGGAAGGTGATTTCAAATTTGAAATTCAAAG
 GAGGAACAAAAGAGAAGACAGAAGCTTCTGGTGGGCAACCTGGAGATTCTGTGAGAGGAGGCTCTTTAAGCTGG
 ATGTTCTCTGAAAGCAGGCGGTGCTTTGTTAAGGTGAGGGCCTACCGGAGTGAAGAGTTCTTGCCTAGTGAGCAG
 ATCCAGGGGGTGTGATCTCCGTGATTAACCTGGAGCCTAGAAGTGGCTTCTTGTCCAACCTAGGGCCTGGGG
 CCGCTTTGACAGTGTATCACAGGCCCCAACGGGGCCTGTGTGCTGCTTCTGTGATGACGAGTCCCCCTGATG
 CCTACTCTGCCTATGTCTTGGCAAGCCTGGCTGGGGAGGAAGTGAAGCAGTGGAGTCTTCTCTAAATTCAC
 CCAATGCAATTGGCGTCCCTCAGCCCTATCTCAACAAGCTCAACTACCGTGGACGGACCATGAGGATCCACG
 GGTTAAAAAAGACAGCTTTTCCAGATTAGCATGGCCAAAGCCCAAGCCCAACTCAGCTGAGGAGAGCAATGGGCCCCA
 TCTATGCCTTTGAGAACCTCCGGGCATGTGAAGAGGCACACCCAGTGCAGCCCACTTCCGGTTCTACAGATT
 GAGGGGGATCGATATGACTACAACACAGTCCCCTCAACGAAGATGACCTATGAGCTGGACTGAGGATCCACT
 GGCATGGTGGCCAAAGCCGATGGAATTCAGGGCCTGTATATCAAGGTGAAGATTGTGGGGCCACTGGAAGTGA
 ATGTGCGATCCGCAACATGGGGGGCACTCATCGGCGGACAGTGGGGAAGCTGTATGGAATCCGAGATGTGAGG
 AGCACTCGGGACAGGGACAGCCCAATGTCTCAGTGCCTGTCTGGAGTTCAAGTGCAGTGGGATGCTCTATGA
 TCAGGACCGTGTGGACCGCACCTGGTGAAGGTCACTCCCCAGGGCAGCTGCCGTGAGCCAGTGTGAACCCCA
 TGCTGCATGAGTACCTGGTCAACCACTTGCCACTTGCACTGACAGGACCCCTGACAGGACCCCTGACAGGAT
 CCCTTGGACCCACTGGGGCCAACTATGGCATCTACACTGTCACTGACAGGACCCCTGACAGGACCCCTGACAGG
 CGGCTCGGCGGCTGCTTTGATGGCACATCCGATGGCTCCTCCAGAATCATGAAGAGCAATGTGGGAGTAGCCC
 TCACCTTCACTGTGTAGAGAGGCAAGTAGGCGCCAGAGTGCCCTCCAGTACCTCCAAAGCACCCAGCCAG
 TCCCCTGCTGAGGCACTGTCCAAGGAAGAGTGCCCTCGAGGAGCAGCAGCGAGCGAGGGGTGGCCAGCG
 CCAGGGTGGAGTGGTGGCCTCTCTGAGATTCTTAGAGTGTCTCAACAGCCCCCTGATCAACTAAGTTTTGTGGT
 ACTTCAACCTCTTCTGCCCTCATTTTATGTGACAGCCATTGTGAGACTGATGCACAACTGTCACTTGGTTAAT
 TTAAGCACTTCTGTTTTGCTGAATTTGCTTGTGTTTCTTCTCATGCCTTTACTTACTTTTGGTCCCATCTGA
 TTGGCACGTGGCCCCCAATGGCACATAAAGCCCCCTTTGTGAACTGTTCTTTAAATGAAACACAAGAAAT
 GGCCACTGGTAAACTCTGCAGCTTCAACTGTACTTCAATTAATGCCATTAATGCAATATACTTCTCTCTT
 TTTGCATGGTTTTGCCACCTCTGCAATAGTGATAATCTGATGTGAAGATCAAATAACCAATATAAAGCATAT
 TTCTTGGCCTTGCTCCACAGGACATAGGCAAGCCTTGATCATAGTTTATATATAAATGGTGGTGAATAAAG
 AAATAAAACACAATACTTTTACTTGAATGTAAATAACTTATTTATTTCTTTGCTAAATTTGGAATTTAGTGC
 ACATTCAAAGTTAAGCTATTAATATAGGGTGTATAGTTTCTTACCAAGTCTGGAAGAAGACATCTCTGGT
 ATCCACAATTAACAGGTTGCTAAGTATTTGTACATTTCCCTTTGCATTGCTTTTGTCTGTAGAAAC
 CCAGTGTAGCCAGGGCAGATGTCAATAATGCATACTCTGTATTTGAAAAA

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAHPADTLESPGEWTTWFNI
DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGS PREGFWCLNREQ
RPGQNC SNYTVRFLCPPGSLRRDTERIWS PWS PWSKCSAACGQTGVQTRTRICLAEMVSLCS
EASEEGQHCMGQDCTACDLTC PMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTK
TPKLLTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTS LKAATIKAEFVRAETPY
MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAG
EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQ NATNSFYVDVGRCPV
KTCAGQQDNGIRCRDAVQNC CGISKTEEREIQCSGYTLPTKVAKECSCQRCTETRSIVRGRV
SAADNGEPMRFGHVYMGNSRVSM TGYKGTFTLHVPQDTERLVLT FVDRLQKFVN TTKVLPFN
KKGSAVFHEIKMLRRKEPIT LEAMETNII PLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKV
KASVTFLDPRNISTATAAQTD LNF INDEGDTFPLRTYGMFSVD FRDEVTSEPLNAGKVKVHL
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL
DVPESRRCFVKVRAYRSE RFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA
CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPR
VKKTAFQISMAKPRPN SAEESNGPIYAFENLRACEEAPP SAAHFRFYQIEGDRYDYNTVPFN
EDDPMSWTE DYLAWWPKPMEFRACYIKVKIVGPLEVNVR SRNMGGTHRRTVGKLYGIRDVRS
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVIPQGSRRASVNPMLHEYLVNHLPLAV
NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA
QQPLIN

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCAC
 AATATCTTAACCTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAG
 ACCTACCTACCCGTACGCATACATACATATGTGTATATATATGTAACTAGACAAAGATCGCAGATCATAAAGC
 AAGCTCTGCTTTAGTTTCCAAGAAGATTACAAAGAATTTAGAGATGTTTGTCAAGATCCCTGTCGATTCATG
 CCCTTTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT
 CAGATTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA
 TCTGAAAGTGAACTCGATCCTCCGGATATTACCTGTGGAGACCTCCTGAGACGTTCTGTGCAATGGGCAATC
 CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCCTGAGCTGATGTTTGATTTT
 GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT
 CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC
 AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA
 GATGCTTTTACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA
 AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTTCGCGCTTTTG
 CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAAACTCAGAGATTTCTTT
 ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC
 ACGTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAATCTCCATGCCACTGTATGTG
 TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG
 AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAATACCTGTATCCC
 CAGTATTTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA
 ACAACGTGCGCTGCCTGTGCCCGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGC
 AGCTGCGGCTCCGACTCTGGCCAGGGCGCGCCCCGCACGGCACCCCGAGCGCTGCTGCTGCTGACCACGCTGCT
 GGGAAACCGCCAGCCCCCTGGTGTTCTAGGTGTACCTCCAGCCACACCCGACGGGCCCTGTGCCGTGGGGAAGCA
 GACACAACCCAAACATTTGCTACTAACATAGGAAACACACACATACAGACACCCCACTCAGACAGTGTAACAA
 CTAAGAAGGCCTAACTGAACCTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTAAATTT
 TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA
 TTGGAAAGGCTGCGACAGCCCCCAACAGGAAAGACAAAAACAAATCAACCGACCTAAAAACATTGGC
 TACTCTAGCGTGGTGCGCCCTAGTACGACTCCGCCAGTGCTGTGGACCAACCAATAGCATTTCTTTGCTGTGAG
 GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC
 CTGTGCTTTAGTGAACGTTGCTCTGTAACCCTCGTTGGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACA
 TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA
 GCACACACCCCACTATACAAGAGTGGCTATAGGAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTT
 ATTTTCTTGAACCTACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAAT
 GTATCTAATTGCAATCAGCAAAGACTGACATTTTATTTTGTCTCTTTCGTTCTGTTTTGTTTCACTGTGCAGA
 GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATATAACAAGTGAATAAGA
 TTCCACCAAAGGACATTCTAAATGTTTTCTTGTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA
 TAAACGATTTTCAGGAATTTGTATTGCAATTTCTTAAGATGAAAGGAACAGCCACCAAGCAGTTTCACTCACT
 TTAAGTATTTCTGTGTGGACTGAGTACATTCAGCTGACGAATTTAGTTCCAGGAAGATGGATTGATGTTCACT
 AGCTTGGACAACCTCTGCAAAATATGAGACTATTTCCACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAA
 AAAAAA

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFD FEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD
TTKKLRDFFTVDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCCGGCTAAGATTGCTGAGGAGGCGG
CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCGTCCGGGCGAGGTGTCTCATGACTT
CTCTTGTGGACC**ATG**TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAAGGGATGGACTGCC
CCTCTCAGCCTCTACTGATTTTTTACCACACCCAAGATTTTTTGAATGGAGGAGACGGCTCA
AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT
AGTATACATTTTTTCTTCTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC
AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA
CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTAG
AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGAAAAAAT
TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA
ATGGGGTGATGAATGGTACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA
CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT
CATTTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT
TGGACCAAACCTCG**TGAG**CCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT
CCGGGAGCAGTGATGTCAAACCTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA
AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTTCGGAAACATCTGCCCATGTGTATTG
ATGGCAGAGCTGTTGCCACAAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT
TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT
GCTTCAGAAGTGTTATTTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT
TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAACCAT
TTACTTTACATATTTCGTTTTCAATACTTGCTGTTTCATGTTACACAAGCTTCTTACGGTTTTT
TTGTAACAATAAATATTTTGAAGTAAATAATGGGTACATTTTAAACAACTCAGTAGTACAACC
TAAACTTGATATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAATG
AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAAAAA
AAAAG

76/330

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFS IHF
SSFGDVACMAICSCQCPAAMAFCFLET LWWEFTASYDTTCIGLASRPYAFLEFDSIIQVKW
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLED TDVANGVMNGHTPMHLEPAPNFRMEPVTA
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

77/330

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTAGAAAAGTGAAGTGGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGGAAAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGA
 CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTG
 CCTGCCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT
 TCTGTGCGAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGGCTT
 CCTCCCCGCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA
TGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC
GAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC
 CCGCTTCAAGAAGCCTGCTGAGTTCAACACAGTGGATGATGAAGATGCCACCGTCAACAAGA
 TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTGCTCCTGCCC
 TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCCTCGGAACTACTACATCCAGTGGCT
 CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCAACCTGTCCCTCA
 TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG
 GGTGTCTGGGCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT
 AGGTATGGTGTGGGTGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT
 ATGACTTTTGGGAGTACTATCTCCCTACCTCTACTCATGCATCTCCTTCCTTGGGGTTCTG
 CTGCTCCTGGTGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT
 AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG
 CAGCCCTGACCCGCAGGATCTGTAATCCTACTTCCTGCTGGCTGCCTTTAGACATGGAGCTG
 CTACACAGACAGGTCTGGCTCTGCAGACACAGAGGGTCTGCTGGAGAAGAGGCGGAAGGC
 TTCAGCCTGGCAACGGAACCTGGGCTACCCCTGGCTATGCTGTGCTTGGTGGTGTGACGG
 GCCTGTCTGTGCTCATTTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG
 CCCCAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG
 TGCCGTCAATTCAGGTTGTACTCATCTTTTACCTAATGGTGTCTCAGTTGTGGGCTTCTATA
 GCTCTCCACTCTTCCGGAGCCTGCGGCCAGATGGCACGACACTGCCATGACGCAGATAATT
 GGGAAGTGTGTCTGTCTCCTGGTCTTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCTGGG
 GCTCACTCGCTTTGACCTGCTGGGTGACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACA
 TTGTGTTCCCTCTACAACGCAGCCTTTGCAGGCCTCACCACACTCTGTCTGGTGAAGACCTTC
 ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT
 CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAG**TGAC**CCTCCAGCTGGGGGTGGGA
 AGGAAAAAACTGGACACTGCCATCTGCTGCCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG
 ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT
 GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAAGTGTGGCCT
 CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCCTGATCCCAAATCTGTTTACACATCA
 ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT
 AGGGTGGGGTAGGGGCAGGGAAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCC
 CTTGCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG
 AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA
 AAAAAA

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNK
IALELCTFTLAIALGAVLLLPSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGVLRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRILLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMPGRMQGTSLGQVSFSKLGSGFAGVIQVVLIFYLMVSSVVGFY
SSPLFRSLRPRWHDAMTQIIIGNCVCLLVLSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY
IVFLYNAAFAGLTTLCLVKTFTAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

FIGURE 80

GGCTGCCGAGGGAAGGCCCCCTTGGGTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

FIGURE 81

GACCGACCTTAAAGAGTGGGAGCAAAGGGAGGACAGAGCCTTTTAAAACGAGGCGGTGGTGC
CTGCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTC
TGTCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGTGCTTGGCGGCGGCGGCTTCCT
CCCCGTTGTCNTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGA
AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGT
GTATTATATCAAACTTCTGTTTGCAAACTGTACATCNTCTGCCACATCTTCCTGACCCGC
TTCAAGAAGCCTGCTGAGTTCACCACAGTGATGATGAAGATGCCACCGTCAACAAGATTGC
GCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCCTTCT
CCATCATCAGCAATGAGGTGCTGCACTCCC

FIGURE 82

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTCATCCTGCTGT
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCAT
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA
GGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA
TACTAAGAGAAACCCAGACAATATCCTTGAGGTTTCAGGAATCTGGAGAGTACAGATGCCAG
GCCCAGGGCTCCCCCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCTTAGGCCTCTC
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA
TGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA
AAA

83/330

FIGURE 83

MLLVVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQTKWYHRYL
GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLD FSSEMGFPHAAQANVELLGSSDLLT

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCGGTGT
 GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG
 GAGTGTGTGGAACAGGACCCGGGACAGAGGAACC**ATG**GCTCCGCAGAACCTGAGCACCTTTT
 GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG
 GGGGTGCCTCGAAGTGCCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAACTAGCCCTGCA
 GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG
 CTGCTTATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA
 GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTTCTTTGGGGATTT
 TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATCCAAGAGGAAGTGATA
 TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT
 AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTGCGCAAGAGAT
 GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT
 GCCCTAATGTCAAACCTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG
 AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG
 AGATTTACGGTTCCGAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATT
 TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT
 CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT
 ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA
 TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA
 CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TG**AGAGTG
 AATAAAATTGGACTTTGTTTTAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTT
 TTGTGTGTGTTTTTGTTTTTATTTTCAATATGCAAGTTAGGCTTAATTTTTTTTATCTAATGA
 TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTTCGAAAAGAATGACC
 AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT
 TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA
 GTTGTTAGCAATTTCAATCAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTG
 TTATTTTAA

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKRKQYDTYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGT PRQQ
DRNIPRGSDIIVDLEVTLEEYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGM EYPFIGEGEPHVDGEPGD LRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGS LIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAGG
CCTATAGGAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCCTGATGATCCACAAGCCCAG
GAGAAATTCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

FIGURE 87

GGCACGAGGCGGCGGGGCAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATGG**
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC
TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG
TACCCTCCGTTGGACCCCCAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTGAG
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC
AGTCTCTGTGCGGCTGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTT**TAGT**GCCT
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG
TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACGTGGCTGGTGAGT
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACCTCTCTGCAAGAGGAGTATTGAAAA
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT
AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG
TGTGGCAATTCTGATCTGCATTTTCAGAAGAGGACAATCAATTGAAACTAAGTAGGGGTTTC
TTCTTTTGGCAAGACTTGTA CTCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAAGTTTGGGTTTGAAGCTGAGGAACT
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAAAAAA

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL
ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTTEKLVAMTMGSGAKMKTSAS
VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLSVSHLVLVTRNACHLTGGLDWI
DQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTCCGAGGTGCTTTCGCCGCTGTCC
CCACCACTGCAGCCATGATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA
TTTGGAGTGTTTTCTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT
TGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTCAGAT
TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTTCAG
GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT
TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTATAACAACA
AGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA
ACTAAGAAGAAGTCAGCAAGCAAAGTGAAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA
ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA
CTGTGGTGCCTGTTTCTTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT
TTTTAGAAGTGCTCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA
TGCATGAATTCGATTGGATTGTGTCATTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG
ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

FIGURE 90

MISLTDQTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK
HKMKATGFFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFI RRV PVLGSLLNLPGI
RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTCTGTCTTTGGA
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT
GGCTTTTGTAATTGGTTTAGAAAGAACATTCAGATTCTTCTTCCAAAACATAAAATGAAAG
CTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTTC

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA
GGCTGCCAGGAAGGAGACGCCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGA
CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC
TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC
ATCAACACCATTTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA
GATCAACTGCAGACTGTCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT
CGGGCACGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAAT
GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA
ACGCTTTGGGCTGTTAGGGGGCTCCAAGGTCCTGGCCAAGAAAGAGCTGGCCTATGTCCCAA
TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGAT
CGCAAGACGGTTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCT
GATTCACTGTGAGGGCACACGGTTACCGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC
GGGCCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC
ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTCAGAAA
TAATGAAAATCCAACACTGCTGGGAGTCCTAAACGGAAAGAAATACCATGCAGATTTGTATG
TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC
AAGCTCTACCAGGAGAAGGATGCCTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCCAGA
GACGCCCATGGTGCCCCCCCCGGCGGCCCTGGACCCTCGTGAACCTGGCTGTTTTGGGCCTCGC
TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG
CTGGCCAGCTTCAICCTCGTCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGT
GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGAC**T**
GACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAACTGGTGGCCTCTGCATATCCT
CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA
CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT
TTTTCCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGC
TGTGTGGTGAGTGTGAACTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG
GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT
AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC
CAAGAAAAAAAATTAAGTGCTTTTCTGGGTCAAAAAAAAAA

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHNKFEIDFLCGWSLSERFGLLGGSKVLAKK
ELAYVPIIGMMWYFTEMVFC SRKWEQDRKTVATSLQHLLRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

93/330

FIGURE 94

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG
 GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA
 GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA
 TACAATTGACATTTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG
 GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT
 TGGTACAAATTCCGTCGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA
 AACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA
 TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCCTTATATAAACCTCAAAAAGGACTT
 TTTCACAGGGTACCTTTAGTGTTGTTGCCAATCTGGGCATGTCTGAACAACCTGGGTTATAAAAC
 TGTATCAGGTTCCCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT
 TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA
 CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT
 AGTAAAGGATGTAAACAGATTAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG
 CAGCAAGAGAGAAGAACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA
 CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAATAGACA
 TGTTTTCTAAAAGTAGCTGTAAC TACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA
 TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAAGCAT
 AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA
 CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC
 CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTTGGTGAATATTCACGGTCTCCTACA
 TTT**TGA**TCCTTTTAACTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT
 TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC
 CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC
 ATCAGATGCTTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCT
 TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA
 TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC
 TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGAAAAGCAAGAATAGCCTTAT
 TTTCAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG
 TGATACTTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA
 ATAAATTTGCAAAACATCATCTAAAATTTAAAAA AAAAAAAAAAAAAAAAAAAAAA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGKTVSGSC
MSTGFSRAVQTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLLDTDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGKTVSGSC
MSTGFSRAVQTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLLDTDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

FIGURE 96

GGCACAGCCGCGCGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGC
 CCAAGCAGCGCGCAGCGAACGCCCGCCGCCACACCCCTCTGCGGTCCCGCGGCGCCTGCCACCCCTTCCCT
 CCTTCCCCGCGTCCCCGCCCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCCGCGAAACCCCGAGGTACCA
 GCCGCGCCTCTGCTTCCCTGGGCGCGCGCCGCCCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACC
 GGGGACCGTTGCCTGACGCGAGGCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCTGCTCGCCTCTTCCAC
 CAACTCCAACTCCTTCTCCCTCCAGCTCCACTCGCTAGTCCCGGACTCCGCCAGCCCTCGGCCCGCTGCCGTAG
 CGCCGCTTCCCGTCCGCTCCCAAAGGTGGGAACCGCTCCGCCCGGCCCGCACCATGCGCACGGTTCCGGCTTGCC
 CGCGCTTCTCTGCACCCCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGG
 AAGTGCGACGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCAT
 TTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA
 TGATTTCAAAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTACGTTACAAGAAGTTTG
 ATGAATTCTTCAAAGAACTACTTGAATATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT
 TTATACATGCAAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT
 GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCTGGTGAACCTCCAGT
 ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT
 CGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTTCGCTCAAGGCTTAGCGGTTGCGGG
 AGATGTCGTGAGCAAGGCTCCTCGTGTTAAACCCACAGCCCAAGTGTACCCATGCCCTGTTGAAGATGATCTACT
 GCTCCCACTGCCGGGTCTCGTGACTGTGAAGCCATGTTACAACCTACTGCTCAAACATCATGAGAGGCTGTTTG
 GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA
 GGGTCTTTCAACATTGAATCGGTTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG
 ATAATAGTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATT
 TCTCGTTCCATCTCTGAAAGTGCCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCCAACACAGC
 AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTCTGGTCTCT
 CCTTCCGAGCAACGTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG
 AAAGGCCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCA
 GGTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA
 AGAATGCATACAATGGGAACGACGTGGACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGT
 GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA
 GAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCTGG
 TTATGCAGAGAGAGTGGAGATTAATTTCTCAAACCTCTGAGAAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAATT
 ATCACTTTTCTACCATCCTAGTGACTTTGCTTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGC
 CACTGGTTTAAAGAGTGCTGACTTTGTTTTCTCATTTCAGTTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGT
 TCCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTTGTGA
 TTTTATCACTCTATTATTTGTTTGTATGTTTTTTCTCATTTCGTTTGTGGGTTTTTTTTTCCAACGTGTGATCT
 CGCCTTGTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAATA
 GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAAAGAAAAAGCCCAAAAAGC

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLYMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFRLVNSQYHFTDEY
LECVSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPPEE
RPTTAAGTSLDRLVTDVKEKLQAKKFWSSSLPSNVCNDERMAAGNGNEDDCWNGKKGKSRYL F
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

FIGURE 98

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA
GCAACTTACAGCTGCACCGACAGTTGCG**ATG**AAAAGTTCTAATCTCTTCCCTCCTCCTGTTGC
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCAC
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA
AGATTGGTTCCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGC
AGTGCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGA
AAGCCAAACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT
TGCTCTGCCTTTG**TAG**GAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACC
CCTAAATCATTCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT
AATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCTAGCTAGTGTCAATTAACCTTAAATGC
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTAAATGTCAAAAAAAAAAAAAAAAAA

101/330

FIGURE 101

MAVLVLRLTVVLGLLVLFLTCYADDKPKDPDDKPDGSGKDPKPDFPKFLSLLGTEIIENAVE
FILRSMRSRSTGFMEFDDNEGKHSSK

101/330

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT
CAGAGCTGGTCTGCC**ATGG**ACATCCTGGTCCCACTCCTGCAGCTGCTGGTGCTGCTTCTTAC
CCTGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAGCTACTTCC
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG
CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC
CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT
GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCC
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCT**TAG**CAGAATGAGAGAAGACATT
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT
CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT
CTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA
AACCACG

FIGURE 103

MDILVPLLQLLVLLLTPLPLHLMALLGCWQPLCKSYFFPYLMAVLTPKSNRKMESKKRELF SQI
KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV
APGEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLEFFWEHVAEPYGSWAFM
WQQVFEPWTKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP
SSKALICSFPSLQLEQATHQPIYLPLRGT

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGGAAGTTGCCTCATCGCAGG
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA
ACTAATATTTATATGACAGAAGAAAAAGATGTCATTCCGTAAAGTAAACATCATCATCTTGG
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG
TTAAGGAATGAGGTTACAGATTGAGGAATTGTAGGGCCCTCAACCTATAGACTTTGTCCCAA
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATCTG
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAAT
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAAACTTTTGGAAGGAA
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC
TTGCCAATTCTGGTTCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG
AAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTGTAAGCTTTCCATGAAAGCCAGCACTTG
CTCATTTAATCCTGGAGTTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA
ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT
GGTAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT
ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTC AACCTAATCCGAAG
ATATACCGAGATCTCAAACATAAAGTGAAACAGAATTTGAACTGTAAGCAAGCATTTCTCAG
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTCCTT
ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA
TAATTCAAAACCTGCTGTTGGTTTTAATTTTGTAACTGTGGCCTGATCTGTAAATAAACTT
ACATTTTTTC

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTD SGIVGPQPIDFVPNALRHAVDGR
QEEIPVVIAASEDRLGGAIAAINSIQHNTRS NVIFYIVTLNNTADHLRSWLNSDSLKSIRYK
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT
ALKPGHAAAFSEDCDSASTKV VIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA
NLTEWKRQNITNQLEKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTIDPMWNVRLGS
SAGKRYSPQFVKA AKLLHWNHGLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

FIGURE 106

TGGTTTTTGCCCCATAAATTCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT
GGGAGACAAGAGGAGATTCCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT
TGCAGCTATAAACAGCATTTCAGCACAACTCGNTCCAATGTGATTTTCTACATTGTTACTC
TCAACAATACAGCAGACCATNTCCGGTCCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCA
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCACAGCG
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTTCAGCCTCTAC
TAAAGTTGTCATCCGTGGAGCAGGAAA

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG
 GAGCTAGAGGGCAAGTGTGCTCGGCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC
 TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCGGGGCGCGGGCTGCA
 TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT
 GTTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT
 TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT
 CCGTAAAGTAAACATCATCATCTTGGTCCTGGGCTGTTGCTCTCTTCTTACTGGTTTTTGCAC
 CATAACTTCCTCAGCTTGAGGCAGTTTGTAAAGGAATGAGGTTACAGATTACAGGAATTGTAG
 GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA
 GGAGATTCCCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA
 ACAGCATTACAGCACAACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA
 GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG
 TCAATTTTGACCCTAAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC
 ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGG
 CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA
 CTGAAGCCAGGACATGCAGCTGCATTTTTCAGAAGATTGTGATTACGCCCTCTACTAAAGTTGT
 CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA
 TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC
 CTGACGGAATGGAACGACAGAATATACTAACCAACTGGAAAAATGGATGAAACTCAATGT
 AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACACACCTCCTCTGCTTATCG
 TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT
 GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA
 TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA
 GACCCAACAGGCAAATTC AACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA
 CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG
 TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG
 GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT
 ATGTCTCCATCTGCCTTACCAAGTGT'TTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA
 CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAACCTGCTGTTGGTTTTAATTTT
 GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTCAATAGGTAAAAA

FIGURE 108

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG
TCTGTGCTGGTCTGAGGGTGCTGCCTGTC**ATG**GGGGCAGCCATCTCCCAGGGGGCCCTCATC
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCATCCTCTGCTG
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC
TGGCCCCTGTCCTGAGAAGGCCCCACCACCCAGAAAGCCCAGCCATGAAGGCAGCTACCTGC
TGCAGCCCTGAAGGCCCTGGCCTAGCCTGGAGCCCAGGACC**TAA**GTCACCTCACCTAGAG
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTTCCCTAGGCTGAGCACC
TCTAGGCCCTCTAGGTTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG
GCTGGGCCCCCTCCCCTGGTCCTCCCAGTGTTTGCTGGATAATAAATGGAACCTATGGCTCTAA
AAAAAAAAAAAAAAAAAAAA

109/330

FIGURE 109

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH
PRSPAMKAATCCSPEGPWPSLEPRT

109/330

FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA
 GTTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA
 CTCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGA
 TCA**ATG**TCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC
 ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCAATTGGTTAT
 TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC
 TCAGCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTA
 TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT
 GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC
 AGCCACTGTGGACATTTGCCATCCTCATTTTTCTTCTGGGTCCCTCTGGGTGGCTGTGCTGCTG
 AGCCTGGGAAGTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCCT
 TTCGGGCATTTCGGTACATGTGGTTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA
 TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT
 AAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCA
 AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA
 TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA
 TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA
 TACTACAAGTGTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT
 TGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAATTTTTCTA
 GGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCG
 GGCATTCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTTGCCTACTTAGTAGCCC
 ATAGTTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGAT
 CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT
 CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA
 ATGAGGAGGGAACAGAACTCCAGGCCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCT
 GGAAAACATTTCCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT
 AGTGAATTTTTTTTTTAAAGACCTAATAAACCCCTATTCTTCCTCAAAA

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

[illegible]

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTPSALTYSSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQMDLWSTATYTTSSQNRPRADPGIQ
RQDPGAAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTQLPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPD
LNAIKRVGYWSSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRI PAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRI IQNICTEFNSQPRN

FIGURE 114

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGGCGCTGAGAGGACACGAGCTCT**A**
TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT
 GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT
 TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG
 TGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGAAACAAACATT
 CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA
 GGCTGGATGGCCCTGTTCCGGGCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACCTCC
 TCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGAACAGTGAACCTACTTCATGGCGTG
 AACCCAGGAGAGACCCCTGTACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC
 CACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCAAGATGTGGCCAGAGTGGCTTTGATGC
 GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCTCACTGGC
 AAGTGGGTGGCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT
 GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCTAGAGTATAACAAAG
 CCATCCGGAACCTACACCCGCTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACT
 GTGTCCATGCCAGTCTTCCAGTCTTGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG
 AGACATTGACAATGCCATGAGGACCTTCCTCACTACTACACTGTATGGAAGCAGTTTGGGG
 GGCTCCCGGAATTCTACAACATTTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA
 CTTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT
 CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT
 TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG
 GCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACCAACTTCATCCACAACAATGG
 GTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGCATCCTGGGGGCTGGGGGGTACA
 TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG
 GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTC
 GAAATTTAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC
 TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCA
 CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCTCT
 AGACTCCTCA**TAA**CCACTGGATAATTTTTTTATTTTTTATTTTTTTGAGGCTAAACTATAATA
 AATTGCTTTTGGCTATCATAAAA

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHL LSKKAGVEVEAGWPCSGPLL RMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLVVQMYKGTVSMPVFQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNI PQGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

FIGURE 116

AAAGTTACATTTTCTCTGGAACCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
 GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
 ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
 GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
 ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
 TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCAGTGATCGCGCCTGGAGA
 AACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
 GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
 ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
 GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
 TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
 CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
 GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
 CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
 GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
 GGTTCGTGCCACTGTTTCGTCTGGAATAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
 TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCAGAAAGTTAATCAGCTGCAGAAGG
 GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
 CTCATAGGTTTGCGBAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
 ATGAGGGGGACAAGTTGTGTTTCTGTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
 GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
 TGA CTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
 CTGGGAAAAGTGA CTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
 TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
 TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTAG
 TGTTCCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
 GGCTTGGAGAGCCCACTTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
 TGTGAGTTCACTTCAAGCCCAATGCCGGTGACAGAGGGGAATGGCTTAGCGAGCTCTACAGT
 AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
 CATGA ACTACTGTAAAGTGTTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
 GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTCT
 TGTGTTGTAAGTACAGAATTACAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
 AAAAAAAAAA

FIGURE 117

MQTFTMVLEEIWTSLEFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAIGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVQGEAIPVLALFAFVGFMILILV
VVPLFVWKMGRLLQYSCCPVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

FIGURE 118

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCTAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 119

CGGACGCGTGGGCCGCCACCTCCGGAACAAGCCATGGTGGCGGCGACGGTGGCAGCGGCGTG
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG
TCAACATCCGGGGCAAACCTGGTGTGCTGCGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG
AATGTGGCCAGCGAGTGCGGCTTACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG
AGACCTGGGGCCCCACCACTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGCACCTACAGTGTCTCATTCCCC
ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA
GACTTCTGGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG
TGGTAGGGGCTTGGGACCCAACCTGTGTGAGTGGAGGAGGTCAGACCCCAGATCACAGCGCTC
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTATATAACCACCGCGTCTCCTCCTCCACCA
CCTCATCCCCGCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAG
AGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA
AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA
TAGAAGTATATCAAGCAATAATCTCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT
AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT
GTTATTTCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA
ACAATACCTCACGATATAAAATAAAAAATGAAAGTATCCTCCTCAAAAA

120/330

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQ
HYRALQQQLQORDLGPHHFNVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG
AHPAFKYLAQTSGKEPTWNFWKYL VAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

120/330

FIGURE 121

CGGACGCGTGGGCGGGCCGGGACGCAGGGCAAAGCGAGCC**ATG**GCTGTCTACGTCGGGGATGC
 TGCGCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGGCCCGGGCCGCCCTCTCT
 CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG
 CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAAGGGGTGCACCAAAAAGCATCTTA
 ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCAGAACGAGAGGCC
 TTGGTCGTCTCCATGAAGACGTCAGGTTGACCTTTGCCAACTCAAGGAGGAGGTGGACAA
 AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC
 CTAACCTCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCAGGCGGGCATCATTCTGGTG
 TCTGTGAACCCAGCCTACCAGGCTATGGAAGTGGAGTATGTCTCAAGAAGGTGGGCTGCAA
 GGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT
 GTCCAGAAGTGGAGAATGCCAGCCAGGGGGCCTTGAAGAGTCAGAGGCTCCAGATCTGACC
 ACAGTCATCTCGGTGGATGCCCTTTGCCGGGGACCTGCTCCTGGATGAAGTGGTGGCGGC
 TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCTGTCTGCCATG
 ACCCATCAACATCCAGTTACCTCGGGGACAACAGGCAGCCCCAAGGGGGGCCACCCTCTCC
 CACTACAACATTGTCAACAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC
 ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCTGTACCATTGCCTGGGTTCCGTGGCAG
 GCACAATGATGTGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC
 AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCTGTATGGTACCCCCACGAT
 GTTCGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG
 GTGTCAATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT
 ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTT
 CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGCGCAGAATTATGCCTCACACGGAGG
 CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGC
 ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT
 GGATCAGGACAAGTGGTATTGGACAGGAGATGTCGCCACAATGAATGAGCAGGGCTTCTGCA
 AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG
 CTCGAGGACTTCTTTCACACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA
 CGATCGGATGGGGGAAGAGATTTGTGCCTGCATTCGGCTGAAGGACGGGGAGGAGACCACGG
 TGGAGGAGATAAAAGCTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC
 GTGTTTGTACAAACTACCCCTCACCATTTAGGAAAGATCCAGAAATTCAAACCTTCGAGA
 GCAGATGGAACGACATCTAAATCT**TGA**ATAAAGCAGCAGGCCTGTCTGGCCGTTGGCTT
 GACTCTCTCCTGTCAGAAATGCAACCTGGCTTTATGCACCTAGATGTCCCAGCACCCAGTTC
 TGAGCCAGGCACATCAAATGTCAAGGAATTGACTGAACGAATAAGAGCTCCTGGATGGGTC
 CGGGAACCTGCCTGGGCACAAGGTGCCAAAAGGCAGGCAGCCTGCCAGGCCCTCCCTCCTG
 TCCATCCCCCACATTCCCCTGTCTGTCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT
 GAAAAAAAAAAAAAAAAA

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRWQEARLQGVFLSSREVD RMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLP GTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCKIVGRSKDMIIRGGENIYPAELEDDFFHTHPKVQEVQVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277,
314-319, 318-323, 379-384, 380-385, 409-413

N-glycosylation Site:

amino acids 282-285

FIGURE 123

CAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGACACCAGAGCAGTTGCGGA
TGATCCTGCCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAGGCACAATGATGTGTCTG
ATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGCAAGAAGGCACTGGAGGC
CATCAGCAGAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGATGTTTCGTGGACATTCTGA
ACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCATTGCTGGGTCC
CCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGT
TGCTTATGGAACCACAGAGAACAGTCCCCTGACATTCGCGCACTTCCCTGAGGACACTGTGG
AGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGGCGCGGATCATGAACATG
GAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGCATCCGAGGGTACTGCGT
CATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGT
ATTGGACAGGAGATGTCGCCAC

FIGURE 124

GAGCAGGACGGAGCCATGGACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC
 AGGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG
 TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC
 GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC
 AGTGCGGGGTTGCGGTTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC
 TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC
 ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGATACCCGCCCAACGGCGTGGAGTG
 CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCT
 GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA
 GCTAATGTGACTGTGTCTTGCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA
 TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAAC
 CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT
 CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAG
 ACCCACATCCACCACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG
 AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC
 CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG
 TGTGGCTCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTTACTGT
GAGCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT
 CATCACTTCCTGTTCCCACCACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC
 CAGTATCCCCAGCTTCTGCTGCGCTGGTTTGCGGCTTTGGGAAATAAAATACCGTTGTATAT
 ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC
 TCCGCTTGTCTCTTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTG
 AGAGAGAGGATGCTAAGCTTCCTACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGG
 GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG
 GAATCGGTTCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC
 CCAATTCGCCCTATAGTGAGTCGTA

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT
EAVGAVETIHGQFSLAVRGCGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL
DPAGNESAYPPNGVECYSCVGLSREACQGTSPFVSCYNASDHVYKGCDFGNVTLTAAANTV
SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTT
VASTTSVTTSTTSAPVRPTSTTKPMPAPTSQTTPRQGVHEASRDEEPRLTGGAAGHQDRSNSG
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT
EAVGAVETIHGQFSLAVRGCGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL
DPAGNESAYPPNGVECYSCVGLSREACQGTSPFVSCYNASDHVYKGCDFGNVTLTAAANTV
SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTT
VASTTSVTTSTTSAPVRPTSTTKPMPAPTSQTTPRQGVHEASRDEEPRLTGGAAGHQDRSNSG
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

FIGURE 126

CGGGACTCGGCGGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC**ATGG**AGT
TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAG
GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTT
CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT
TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCAGGTGGAGAACCTCATCACCGCC
AATGCAACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAA
CCTGAGGCGGCTGCTTGAACCTTTGGATGCAATGTGATGCT**TAA**GAAAACCGGCCACTTC
AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCACCCTATCCCCTCTA
ACACCATTCCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCCT
GCCACCTCCCGTGATGTGTGTGTGTGTGTGTGTGTGTGACTGTGTGTGTTTGCTAACTGTG
GTCTTTGTGGCTACTTGTTTGTGGATGGTATTGTGTTTGTAGTGAAGTGTGGACTCGCTTT
CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCGTGGCCCTCCATCAC
CTTCTGCTCCTAGGAGGCTGCTTGTTGCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC
GTAGGGTAAGAGCACGGGCAGTGGTCTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCAC
TTTGTGCATCATTCCTCATGGACTCCTTTACTCCTTTAACAAAAACCTTGCTTCCTTATCCC
ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC
CAGCGTTGACGTCAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG
AGGAGTCCCCATCTGCCCCGCCCCCTTACAGAGCGCCCCGGGGATTCCAGGCCCAGGGCTTCT
ACTCTGCCCCCTGGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG
GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG
ATGCAGACTACAGTCCCTGCAATTGGGTCTCTGGCAGGCAATAGTTGAAGGACTCCTGTTCC
GTTGGGGCCAGCACACCGGGATGGATGGAGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACG
TCCCCTTAGATGGGCAGCAGAGGCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGA
GCGGTGAGCGAGGTGGGTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGAACAGTGAGAG
GTTGAAGGTCATAACGAGAGTGGGAACCTCAACCCAGATCCCGCCCCCTCCTGTCTCTGTGTT
CCCGCGGAAACCAACCAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT
CCTCAACAACAACAGAAAAAGGAATAAAATATCCTTTGTTTCCT

127/330

FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK
CSFNQKPRAPGDEEAQVENLITANATEPQKQRTQVQPSGGSLWNLRRLLLEPLDANVDA

127/330

FIGURE 128

AAACTTGACGCCATGAAAGATCCCGGTCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCTCTGCCCAGGGAGCCACCCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG
CGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCTGCGTTT
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAGGAACTTCCTTT
CCTCAACTGGGATGCCTTTCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT
CTTTCCTGCCTCAGGAAGTCCAATAAAACATTTTCCATCCAAA

AAACTTGACGCCATGAAGATCCCGGTCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCTCTGCCCAGGGAGCCACCCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG
CGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCTGCGTTT
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAGGAACTTCCTTT
CCTCAACTGGGATGCCTTTCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT
CTTTCCTGCCTCAGGAAGTCCAATAAAACATTTTCCATCCAAA

FIGURE 130

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCC**ATG**GGGGTGGAGATTGC
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTC
TTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC
TTTGCCGGCCACTCA**TGA**GAGTGTTTTTGTGTAAAGTATTTTTTTAGAATACTGTTGACTTCT
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT
AGGTTCAAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTTTATC
TTCATATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT
TATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAA
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT
AAAAGTTTTGAGTATATTTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

131/330

FIGURE 131

MGVEIAFASVILTCLSLLAAGVSQVVLLQPVPTQETGPKAMGDLSCGFAGHS

131/330
MGVEIAFASVILTCLSLLAAGVSQVVLLQPVPTQETGPKAMGDLSCGFAGHS

FIGURE 132

GGGGAATCTG CAGTAGGTCTGCCGGCGATGGAGTGGTGGGCTAGCTCGCCGCTTCGGGCTCTG
GCTGCTGTTGTTCTCTGCCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAATGGA
AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAC
TGCAGCTGCTACCATGGTGT CATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG
GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA
GACTGTACCGGGAAAATGACTGCATGTTCCCTCAAGGTGTAGTGGTGTGAGCACTTTATT
TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA
GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC
ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT
CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG
GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG
ATCCTCTCATTCCTTCTGTCTCGGAAAAACCCAAAACCTTGTTGATGCAGAATACACCAAAAAC
CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT
GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA
AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC
TATCCACAGCTGAAGCCATGGGTTCACTATATCCAGTCAAACAGATCTCTCCAATGTCCA
AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA
GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG
AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT
TCCCAAAATGTTGAAAAC TGAAC TA TAGT CATCATAGGACCATAGTCCTCTTTGTGGCA
ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA
TCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGA
AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA
CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC
ACAGCTTGTCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA
TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA
CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG
AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG
TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTA AACCATAACTCTGTTACTCAG
GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT
TCAGGTTCCCTTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

FIGURE 133

MEWWASSPLRLWLLLFLLP
SAQGRQKESGSKWKVFIDQ
INRSLENYEPCSSQNCSCY
HGVIE
EDLTPFRGGISRKMMAEV
VRRKLGTHYQITKNRLYRE
NDCMFPSRCSGVEHFILEV
IGRLPD
MEMVINVRDYPQVPKWME
PAIPVFSFSKTSEYHDIMY
PAWTFWEGGPAVWPIYPT
GLGRWDL
FREDLVRSAQWPWKKNST
AYFRGSRTSPERDPLILLS
SRKNPKLVDAEYTKNQAWK
SMKDT
LGKPAAKDVHLVDHCKYK
YLFNFRGVAASFRFKHLF
LCGSLVFHVGDEWLEFFYP
QLKPWVH
YIPVKTDLSNVQELLQFV
KANDDVAQEIAERGSQFIR
NHLQMDDITCYWENLLSEY
SKFLSY
NVTRRKGVDQIIPKMLK
TEL

FIGURE 134

CACCCCTCCATTTCTCGCC**ATG**GGCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT
TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT
GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA
CCGCAGCATCCTTGCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC
ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG
AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCACTGGTGGTACTGGGAGCC
CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC
TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT
GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC
TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCACCTGCGCCACCCAGTGTGTG
TGGAGCTGCTGACAGTGTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT
TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT
CCGGGCCCAGCTACAAAGAAACTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAGT**T**
GAGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCCTCTCCCACTGAATTCTAAATCCTTAAC
ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCTT
CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA
CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC
TTCCTGTTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC
CTGACCACTCCCCTGGCACTGTTACTTGCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT
GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA
GGCCCCAACCTTGCCTCACCCTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT
GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACCTCGATGACTTGGGGCTC
CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTCTAAAAAAA

FIGURE 135

MAPALLLI PAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAAALQDRSILAP
LAWDLGLLLLFVGQHSLMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP
RALRLFSLRHPVCVELLTVLWVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR
KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTT TAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA
 AGAAATTGCCAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTT CACAACAGATCTGAGTGT
 TTTAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTT CATCTGGAATT
 CCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTA
 TCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTAAGAACCTTCGGCTCGCGTGCTTCTG
 AGCTGCTGTGGATGGCCTCGGCTCTCTGGA CTGTCTTCCGAGTAGGATGTCACTGAGATCC
 CTAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT
 TCCCCACTACAATGTGATAGAACGCGTGA ACTGGATGTACTTCTATGAGTATGAGCCGATTT
 ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAA ACTGCTCTCATCAAAATCCATTT
 CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTAC
 TTGGGGTGAAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG
 AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT
 GACATAATCCGACAAGATTTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTTATGGC
 ATTCAGGTGGGTAACTGAGTTTTGCCCCAATGCCAAGTACGTAATGAAGACAGACACTGATG
 TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT
 TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT
 TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATATAA
 TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGT CACGTAAAACCCATCAAGTTT
 GAAGATGTTTATGTCTGGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGA
 CACAAATCTTTTCTTTCTATATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTG
 CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTTGGCAGGTCATGCTAAGGAACACC
 ACATGCCATTATTAACCTTCACATTCTACAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA
 GTGTTAAATAAAGTAGGTACTGTGGAAAATT CATGGGGAGGTCAGTGTGCTGGCTTACACTG
 AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC
 AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA
 GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG
 TGTTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAAAACAATGTAGAGTTTTTATTTATTG
 AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA
 TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT
 TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTT CACAGTTATTATTATTAAAATTA
 CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT
 CATTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC
 TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGACTTTGTAAAT
 ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPiYRQD
FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EK
EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFIN
TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD
LVPRIYEMMGHV KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG
FSSKEIITFWQVMLRNTTCHY

FIGURE 138

CCTCTGTCCACTGCTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCATCTGGGATTATGGAAATGGCTTTGCTGCAACC
AGACTCTTTCAAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA
AAAAACATTGCAAACATGTGTCGTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT
CCTTCTGTGGAGACACGGTGGAGAACTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAAA

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNGWDSWNS
IWDYGNNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

139/330

FIGURE 140

CATTCTGAACTAATCGTGTGAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATA
 TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG
 CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT
 TGAGTTAAGAGTTGTTATATGCTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAA
 GAATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGCTGGTTTGTTTACATG
 CAAGCTTATAGTTGAAATATTTTTTCAGGAATTAC**ATGA**ATGACAGTCTTCGAACCAATGTGT
 TTGTTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT
 CAGATTCCGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAAT
 CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAGCCAACTATGAATTAC
 TGGAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCAAATTTAAAGCAAAGGGA
 TTGAATCCGGATGGAAGCTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC
 ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG
 TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAAGAAA
 GACAGCAAGAGAAGTAGAAATAGCAGAAGTGCAAGTCGATCGAGGTCAAGAACACGATCAG
 TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA
 GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT
 GGTCTCTCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGACATGG
 TCATAAAAGGAAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG
 CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT
 GAGAGGTCCCATAAAGCAAGCACCATGGTGGCAGTCGCTCAGGACATGGCAGGCACAGGCG
CTGACTTTCTCTTCTTTGAGCCTGCATCAGTTCTTGGTTTTGCCTATCTACAGTGTGATGT
 ATGGACTCAATCAAAAACATTAAACGCAAAGTATAGGATTTGATTTCTTGAAACCCTCTA
 GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTGCACATT
 AAAATGCCCTAGCAGTATCTAATTAATAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT
 TGTGTATTGTTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT
 ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT
 ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTACAAGGAAATAAAATACAAAT
 CTTGTTTTTTCTAAAAAAGAAAAAAGT

FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRRKRSRSQ
SKSRDHSDAAKKHRHERGHRDRRERSRSFERSHKSKHHGGSRSRSGHGRHR

141/330

TGGGGATAAAGGAAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA
TTTTTTGAANNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNNTATTGCTTACTGATTTTTTTG
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAAGA
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTTTGTTTACATGCA
AGCTTATAGTTGAAATATTTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT
GTTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAACTATGAATTACTG
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT
GAATCCGGATGGAACTCCAGCCCTTTCAACCCTGGGGTGGATTTTCTCC

FIGURE 143

GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTC**ATGGC**
GGTCCTCGGAGTACAGCTGGTGGTGAACCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCGGATACAAGCACCCG
TCTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG
AGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCCTGCGCTTCTTCCTGGAGTACCAGTGG
TTTGTGGACTTTGCTGTGTACTCGGGCGCGTGTACCTCTTCACAGAGGCCTACTACTACAT
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT
CCATCAAGATGTTCCCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG
GGAGGAGACCCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCAGGCCTGCGGCTGGC
CCAGACCCACCGGGACGCACTGACCATGTGCGAGGACAGACCCATGCTGCAGTTCCTCCTGC
ACACCAGCTTCCTGTCTCCCTGTTCATCCTGTGGCTCTGGACAAAGCCATTGCACGGGAC
TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA
CTCTGGGCGCCTCTGGTTGCTGGTGGTGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT
GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCCCTCAACTGCACACTTCTGCTCAAGACGC
TGGGAGGCTATTCTGGGGCCTGGGCCAGCTCCTCTACTATCCCCGACCCATCCTCAGCC
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG
GGCCCTGGGTGGCCTGCTTACTCCCCCTCTTCCTCCGTGGCGTCTGGCCTACCTCATCTGGT
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA
GGCTCC**TAG**CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA
CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCAGCTGCAAGGTGGGGCCGGAATCCCC
GGCGTTCCCTTCACCACAGTGCCTGACCCGCGGCCCCCCTTGGACGCCGAGTTTCTGCCTCA
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG
TGCCAGGTTTGAGTGCGGAGGGTGTGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC
GATTTTTTAA

FIGURE 144

MAVLGVQLVVITLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEEELRALAGKPRPRGRKE
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLEPGLASMTQNLEPLLKKQGDWALPVAKLAI RVGLAVVGSVLGAFLTFPGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVSLOQYLTPLIILTNLCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAAARI
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLA GS

FIGURE 145

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTCCCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

GGTCTCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGGCCCGTGATTTTATTAACGTGGCTT
AATCTGAAGTTTCTCAGTCAAATTCCTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGC
TTGGCTGTTTTGGGCCCTTGTAAGCTGACAGAAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGG
CGCTTCTGTTGCTGGTCTTGCTTGGCTCAGTCTCTGCTAACTACATATGCAATGTGGGCAACCTGCACCTTCTCTG
TATTCAGAATCTGTAAAGGTGCCTCCCCTCAGTCTGACCAAGATAGGAAGAGCGCTCACAAGATGGCTG
TCCAGACGGCTGTGCGAGCCTCAGAGCCACGGCTCCCTCCCAAGAGTTTTCTGCAGCTGCCACCATCTCCTTAA
TGACAGACGAGCCTGGCCTAGACAACCTGCCTACGTGTCTTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCA
GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA
AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGG
ACGAGGGAAAAATCTGAAACACCACTGCCCTGAAGTCTTTCCAAAGGTTATTACCACCTGATTCAGATGGTGAA
ATTACAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTTATAGGCTGGTGGGAGGTAGCGAAAC
CCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAG
GAGACATCATTTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCCTGCGG
CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGCCCC
GGATGCTTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAA
TAAAACTGGTGCGCAAGGTGGATGAGCTTGGGTTTTATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA
CATGGTCACTTGAGGAATGACCGTGCTGATTGCCATCAATGGACATGATCTTCGATATGGCAGCCAGAAAG
TGCGGGCTCATCTGATTGAGGCCAGTGAAAGACGCTGTTACCTCGTCTGCTCCCGAGGTTTCGGCAGCGGAGCC
CTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT
CCCAAGCCCTCCATCCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCT
CGGCATGACCGTTCGACGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTATCAGTGTTGAGCCCG
GAGGAGTCATAAGCAGAGATGGAAGAATAAAAACAGGTGACATTTTGTGAATGTGGATGGGCTCGAATGACA
GAGGTGAGCCGGAGTGAGGCATGGCATTTATGAAAAGAACATCATCTCGATAGTAGTACTCAAAGCTTTGGAAGT
CAAAAGATATGAGCCCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCZ
GTGACTGGTCCCCATCCTGGGTGATGTGGCTGGAATTACCACGGTGCTTGATAACTGTAAAGATATTGTATTA
CGAAGAAACACAGCTGGAAGTCTGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAAACCTTT
TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG
CTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTCTGTCGAAGACTGCTGAAAGAACTTAAAGGAAGA
ATTACTCTAACTATTGTTTTCTGGCCTGGCACTTTTTTATAGAAATCAATGATGGGTGAGGAAAAACAGAAAAA
TCACAAATAGGCTAAGAAGTTGAAACACTATATTTATCTTGTGAGTTTTTATATTTAAAGAAAGAATACATTGT
AAAAATGTCAGGAAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAAATTA
AAACTACTAGTTTTTTTTCTAGTGTGGAGGATTTCTATTACTCTACAACATTGTTTTATATTTTTTCTATTCAAT
AAAAAGCCCTAAAACACTAAAAATGATTGATTTGTATACCCCACTGAATTCAGCTGATTTAAATTTAAATTTT
GGTATATGCTGAAGTCTGCCAAGGGTACATTATGCCATTTTTTAATTTACAGCTAAAATATTTTTTAAATGCA
TTGCTGAGAAACGTTGCTTTTCATCAACAAGAATAAATATTTTTTCAGAGTTAAA

147/330

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFFERSTIRSRS
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP
SESLSIRLVGGSETPLVHII IQHIYRDGVIARDGRLLPGDI ILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKVDPEGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQRS
PDI FQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWS PSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRC GDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

147/330

FIGURE 148

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAA**ATG**GTGCTGACCATCT
TTGGGATACAATCTCATGGATACGAGGTTTTTAAACATCATCAGCCCAAGCAACAATGGTGGC
AATGTTTCAGGAGACAGTGACAATTGATAATGAAAAAATACCGCCATCGTTAACATCCATGC
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG
GGTCAAGTACAACCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCAC
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACAT
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG
TGCAGACATTCATGTT**TAG**GATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC
TTGGTTTACTCAAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC
AGTCAGAAAATATAAATGCTGTATTTATA

149/330

FIGURE 149

MKILVAFLVVLTIIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNQLQWYIYEKQALDNMFSNKYTWVKYNPLE
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

149/330

FIGURE 150

GGCACGAGCCAGGAAGTAGGAGGTTCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC
ATGGGGCTCCCTGGGCTGTTCTGCTTGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCGGGCCAACCTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG
GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAG**AATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVPKGRWVLITCCAPQPPPPITYSL
CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK
PVSELNANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDGQVHLQQRPCHRQPANFSFLP
SQTSDWFWCQAANNANVQHSALTVPVPPGGDQKMEDWQGPLESPIALPLYRSTRRLSEEEFG
GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

151/330

FIGURE 152

GGTCCTTAATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG
CTGTCCGGCTGGTCCC GGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT
TACAGAGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCCTGC
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCTTCCATT
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC
CTGGAGCCAAGTGCAGGAGCACCCTCGCCATGTCTCAGGCACAACCCAACCTCAGGGCCAC
AGCCACCACCCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA
TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCAG
GTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTT
TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC
TCTTTTGTGGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT
GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG
AATTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT
TGTTCTGTACTGATATTTAAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 153

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKERPGPRWCAVQGQVDEKTFL
HYDCGNKTVTPVSPGLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:**Signal peptide:**

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

FIGURE 154

GGGAAAGCCATTTGCGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTG
CCTTGGGCCTCACAAATTTTCATTCTGTTTTCTGACTTTCAAGTTATATAACCGTGGAA**ATG**GAG
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAAATCATGAACTCTGGCTGTTGATCAAAA
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAAGAAGCTAGCAGAAGAC
TCAACCTGGCCTCCCATAAACAGGACAGATTATTAGGTGATGGCAAAAATGGATTCTACAT
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAAACCTCAAATTGGGAGGCCAAC
CCACAGAACAGCATTTCTGGGCCAGGCTG**TAA**TCAGAATTGTCGTCGTACATGCTCAACAGC
ATTGCTTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCTTACCTTT
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTTCTGTAT
CATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLA
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGQPTEQHFWARL

155/330

GTTCTCCTTTCCGAGCCAAAATCCCAGGCATGGTGAATTTATGAACGTGCCACACCC**ATGAAG**
 CTCTTGTGGCAGGTAACGTGTGCACCACACACCTGGAATGCCATCCTGCTCCCGTT**CGTCTA**
 CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGGCCCC
 AGAACTGCCCTCCGTTTGTCTCGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCCGG
 GGCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA
 CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCCTGCAGT
 TGGGCAGGAACCTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC
 ACCCTGGAGCTGTTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC
 CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCATCGAAAGCATCCCCCTCTTACGCCTTCA
 ACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT
 GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA
 AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTGAGGGAACCACT
 TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTATG
 AACTCACAGGTCAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAC
 CAACTTGGCCCAATAACCTCTCTTCTTTGCCCATGACCTCTTTACCCCGCTGAGGTACC
 TGGTGGAGTTGCATCTACACCACAACCTTTGGAAGTGTGATTGTGACATTCTGTGGCTAGCC
 TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTGATGCTCCCAT
 GCACATGCGAGGCCGTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT
 TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAAGTTAAGTGT
 CGGACTCCCCCTATGTCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC
 CTCCCGCCACCCAAGGATCTCTGTCTCAACGACGGCACCTTGAAGTTTCCCACGTGCTGC
 TTTACAGACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG
 GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAAGTACAGCTTCTTACCACAGT
 AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCTTA
 CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC
 CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT
 GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG
 CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC
 ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCAGCAGCAACATCCGC
 AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCACAATTC
 ATGACCATATTAAGTACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC
 CTGGGGAACTCTCTGCACCCACAGTCAACACTATCTCTGAACCTTATATAATTACAGACCA
 TACCAAGGACAAGGTACAGGAACTCAAATA**TGA**ACTCCCCTCCCCCAAAAAAATTATAAAAT
 GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA
 TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAGTAAAGT
 CAAAAAGTCAAAACA

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVCT
RRGLSEVPQGI PSNTRYLNL MENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGFLFNLKYLN LGMCNIKDMPNLTPLVGLEELEM SGNHFPEIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHPWNCD CDILW
LAWWLREYIPTNSTCCGRCHAPMHRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFT TTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTT DKMQTSLDEV MKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS
TVTAARTVEIIQVDEDI PAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNLHPTVTTISEPYIIQTHTKDKVQETQI

FIGURE 158

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCC
 TTCTCATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATG
 TGTGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGAAAGAGGGAAGTCGTGGGTTATACCAT
 CCCTTGCTGCAGGAATGAGGAGAATGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAACT
 GCAAGAGCTGCCGAAATGGCTCATGGGGGGGTACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAG
 TGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT
 GTTGGAAAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTATGCTAAACCTGGGTTTGTCTATCCAACATA
 GATTTGTCTATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTTCGTGATGGAGACAAC
 CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAGGATCCTCACT
 CCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGCAT
 GCTCCTCATCCCCTTGTTTTCCATGACGGCACGTGCGTCCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG
 GCAGGCTATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAAAC'TGCTCAGACCC'TGGGGGCCAGTCAA
 TGGGTACCAGAAAATAACAGGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTGTCTT
 TCTTTTGTAACTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGG
 AAACAGCCCATCTGCATAAAAGCCTGCCGAGAACCAGAAAGATTTAGACCTGGTGAGAAGGAGAGTTCTTCCGAT
 GCAGGTTCACTCAAGGGAGACACCATTACACCAGCTATACTCAGCGGCCCTTCAGCAAGCAGAACTGCAGAGTG
 CCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATACCCAGCTCCAG
 TATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAAGTGGAG
 TGGGCGGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGACCCAAGGGTTGC
 GCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGG
 TTCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGG
 GAAGGTCAACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATGACCGGGATG
 AGAAGACCATCCAGAGCCTACAGATTCTGCTATCATTCTGCATCCCAACTATGACCCCATCTGCTTGATGCT
 GACATCGCCATCCTGAAGCTCCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAG
 TCGGGATCTCAGCACTTCCTTCCAGGAGTCCCACATCACTGTGGCTGGCTGGAATGTCTGCGCAGACGTGAGGA
 GCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGGTGGTCACTGTGGTGGACTCGCTGCTGTGTGAGGAGCAG
 CATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTTCTGTGCCAGCTGCGAACCCTGCCCCCTTC
 TGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGGACGAGCATCTCCTGAGCCACGCT
 GGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCTTACCAAG
 GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAATTGAACCATGCTCATGCACTCCTTGAGAAGTGTTC
 TGTATATCCGTCTGTACGTGTGTCAATTGCGTGAAGCAGTGTGGGCCTGAAGTGTGATTTGGCCTGTGAACCTGG
 CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT
 GCCGCTCCACTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAAGAAGACC
 ATATACAAAACCTCTCCACTCCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTG
 ACCAGGGAAGATCTGGGCTTCATGAGGCCCTTTTGGAGCTCTCAAGTTCAGAGAGCTGCCTGTGGGACAGCC
 CAGGGCAGCAGAGCTGGGATGTGGTGCATGCCTTTGTGTACATGGCCACAGTACAGTCTGGTCCCTTTCTTCC
 CCATCTCTTGTACACATTTTAATAAAATAAGGGTTGGCTTCTGAACTACAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

91 8 8 1 00

91 8 8 1 00

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTTCGCAATCAGATTGGAAAAGCTCAACTTGA
AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAAC**ATG**GGC
TTCAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT
GACAGTGGTTGGGTGGGCCACCAGTAACCTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAG
CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACT
AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACCTGTCCTTCTGTGTCTCCTTACCTCAG
AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC
CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC
CTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT
CCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT
TTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC
TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA
GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG
GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTTCAAGGTGAATGGATTCTCT
AACAACTACTGGGGATGGGGAGGCCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG
AATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG
ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACGAGTCTGG
AGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA
TATCAACATCACAGTGGATTTCTGGTTTGGTGCAT**TGA**CCCTGGATCTTTTGGTGATGTTTGG
AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA
AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCTTTTTGTATTTTCTTAGCAGAGCT
CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG
AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT
AAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATAT
TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT
CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCAATTTATCCTGTACAATCATCT
GTGAAGTGGTGGTGTGTCAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA
GGACACAGTGAACCTTGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAG
CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT
TCCAGTGATGCCACCAGAGAATACATTCTCTATTAGTTTTTTAAAGAGTTTTTTGTAAAATGA
TTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAATAATA
TGTCTATCAAATACCTCTGTAGTAAATGTGAAAAAGCAAAA

161/330

FIGURE 161

MGFNLT FHL SYKFRL LLLLL TLCLTVVGWATS NYFVGAIQE I PKAKEFMAN FHKTLILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAE GKKFNRAKLLNVGYLEALKEEN
WDCFI FHDVDLVPENDFNLYKCEEHPKHLVVG RNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWG WGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRTDGLSSCSYKLV SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSYLTCTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTNLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

FIGURE 164

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACTGAGCTCCCAGATCT
GGGCCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCA
CAACAGACGGGACAACCTTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTG
GATGCCCATGTTCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCT
GCGGCTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAG**AACCTACCTGCCCTG
CCCCGTCCCCTCCCTTCCTTATTTATTCTGCTGCCCCAGAACATAGGTCTTGGAATAAAA
TGGCTGGTTCTTTTGTTCCTTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

164/330

165/330

FIGURE 165

MALSSQIWAACLLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH
FPICIFCCGCCHRSKCGMCCKT

165/330

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC
 CTGGATCTTCCACC**ATG**TTCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC
 ATCTCCCTGACTGTCCTCTTACCCCTCTTCTCGTTTTTCATCATAGTGCCAGCCATTTTTTGG
 AGTCTCCTTTGGTATCCGCAAACCTCTACATGAAAAGTCTGTAAAAATCTTTCGTGGGCTA
 CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC
 GGAATCATTGCAAAGGATCCCACTTCACTAGAAGAAGAGATCAAAGAGATTTCGTGGAAGTGG
 TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCC
 GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG
 GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC
 GGTCTGTGGGGGTTAGGAGTGCTGATTCTGGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC
 TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACCTGTGGTGGGATACTTGCCAAAT
 GGGAGGTTTAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGCG
 AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT
 GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC
 ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTGAGAGGCCATGGTGAAGGCCTG
 CCCACACGTCTGGTTTGGAGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA
 CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCAGAAGGAACCTGCATC
 AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCC
 TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA
 TGGTGACGTACCTGCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG
 CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC
 CATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCTGAAGAGGGAGAAGG
 TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC
 AAGGACAGGAGCCGCTCC**TGA**GCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA
 CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCACTGCTGTGTCTTTCCAGACTCCAGGG
 CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT
 GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA
 CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACCTCCCA
 CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT
 GCGGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCGGCCACCCG
 CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC
 CTTGGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTCAGCAAAATG
 AAGGGTTAGATTTTTTATGCTGCTGCTGATGGGGTTACTAAAGGGAGGGGAAGAGGCCAGGTG
 GGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTTCGTACTCCAGGCTAACCCTGAACTCCCC
 ATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG
 GAATGGTGGTGATTCTACCTCACAGGGCTGTTGTGGGGATTAAAGTGCTGCGGGTGAGTGA
 AGGACACATCACGTTCAAGTACAGGCCACAAAACGGGGCACGGCAGGCCTGAG
 CTCAGAGCTGCTGCACTGGGCTTTGGATTTGTTCTTGTGAGTAAATAAACTGGCTGGTGAA
 TGA

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIIVPAIFGVSEFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRSKSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAEELSWNLLSRTNYNFYIISLRLTVLWGLGVLIKYCFLLPLRIALAF
TGISLLVVGTTVVGYPNGRFKEFMKSHVHLMCYRICVRALTAIITYHDRENRPNGGICVAN
H TSPIDVILASDGYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL
LRMMTSWAIVCSVWYLPMPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVDTF
KEEQQKLYSKMIVGNHKDRSRS

167/330

FIGURE 168

GCCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA
CGTCTCCTCCAGGGATGGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC
ACCTGGCAGGCCCAGGCTGTTCCCAACCATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGA
CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG
AAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCAGGAGACCTGGGAGGAC
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT
CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG
GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCG
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTG
GAAGACTCTGCTCTTGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCCCTGAAAGTCCA
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG
CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCCAGG
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACCTCTGCTATGTGATGGGGACTTCCT
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTTGCAATGTGGAGACA
TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH
ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL
YMRHFPPFKALHFYLRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS
LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

FIGURE 170

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCA
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG
GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC
TATTGTCTGGACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGCACTATCA
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG
CTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT
GCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT
GAAGGTGCTGCTGATGACCCAGATTCTTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCT
CCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAAACTCCTAACATATGCCCCCAT
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA
TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATCTTAGACAGCAGTG
CACTCCCCTAAGTCTCTGCTCA

171/330

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSPKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEEKKRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

171/330

FIGURE 172

CTGGTTCCCCAACATGCCTCACCCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC
TCTGGACCCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC
AGCCAGAAGGGGGCAGCTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA
GATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGT
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCT
GTGTCTCCTGTTGGTGCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGGAA
ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA
TAGAACAAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG
AATGTTATCTAGACAGCAGTGCACCTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAA

GAAAGACGTGGTCCCTGACAGACAGACAATCCTATTCCCTACCAAA**ATGA**AAGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAA
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAA
TTTTTCATGAAATTATTTCCTCTTCCTGTTCAATAAATGATTACCCCTTGCACTTAA

174/330

FIGURE 174

MKMLLLLCLGLTLVCVHAEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNERLFLEQ
IHVLENSLVLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI
NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

174/330

FIGURE 175

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAA
TGGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTC
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAG
AAAAAGAGCGTGCTGCAACAACAGAAGTGAATGTTTCTTTCATCATTTTTTCAGTGTGATCA
CAGTCATTGGTGTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTC
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA
CATTCATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCAC
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCT
TGTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTC
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTA
TCAGTAGTTTGAAAAAAAAAAAA

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA
IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA
NCEFSCLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGAGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAATAGTGAA
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC
CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

177/330

178/330

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVN¹LQVAKLNPPPEALAAKLEV
KHCTDQISFKKRLSLKKS²WWK

178/330

FIGURE 179

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC
GAACTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCAC
TCATCACTCCAGGCTCTGCCACTACTTGCTTGAGACACAGGACTGGCCTCCAGGGATGGCCTGA
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAAACTCG
CCCCACCACCCCCTCA

180/330

FIGURE 180

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK
YKSSQKQHSPVPEKAIPITPGSATTC

180/330

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATGT**CGCTGCTGAGCCTGCCC
 TGGCTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC
 CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCC
 AGTGTTTCCACAGCCCCCAAACGGAACTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT
 ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTGGCCACCTATTCCCAGGGCTTTACGGT
 ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA
 CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGG
 CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC
 GCCCGCCTTCCATTTCAACATCCTGAAGTCTATATAACGATCTTCAACAAGAGTGCAAACA
 TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG
 CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG
 TCAGGAGAGGCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA
 GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGC
 TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG
 CACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG
 ATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT
 ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC
 CTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC
 AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCAGCTGCCC
 TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATCTCCCG
 ATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC
 TCATCGATATTATAGGGGTCCATCACAACCCAACCTGTGTGGCCGGATCCTGAGGTCTACGAC
 CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCCTTTCTC
 CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG
 CGTTGATGCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAA
 TTGATCATGCGCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA
 G**TGA**CTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGEIGILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLLASEGS
SRLDMEFHHISLMTLDSLQKCFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLDLDFIDVLLLSKDEG
KALSDEDIKAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

FIGURE 183

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATC**ATG**TATAAGCTGGCCTCCTGC
TGTTTGCTTTTACAGGATTCTTAAATCCTCTCTTATCTCTTCTCTCCTTGACTCCAGGGA
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG
AAAGCAGACTCAAGTACCAACATTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATACA
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGTCT**TGA**AGTGAAATAAGCATCTGT
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAC
AGTGTGGAGAAAAACTAGGCAAACCTACACCCTGTTCAATTGTTACCTGGAAAATAAATCCTCT
ATGTTTTGCACAAAAAAAAAAAAAAAAA

184/330

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDREISFQLSAPHEDARLTPEELERASLLQILPEMLGA
ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

184/330

FIGURE 185

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT
GGGGTTGCTGGTTTAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA
CCACCTCCGCCAGGAACTGCAGGCCACCTGTCTGCAACCCAGCTGAGGCC**ATG**CCCTCCCC
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGAAGATGGAGGTCAAGCAGA
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT
CAGGGGTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG
GAAGAGGCCAAAGAGGCCCCAGCCGACAAG**TGA**TCGCCCACAAGCCTTACTCACCTCTCTCT
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA
CAAGCTCAGGAGGCGAATAAATGTTCAAAGTGTGTA

186/330

FIGURE 186

MPSPGTVCSELLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

186/330

FIGURE 187

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC
 GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAAT**ATGA**ACACGTGGCTGCTGT
 TCCTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG
 CTCCTGCTGGACTTTCTTGCTTGCTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT
 GAGTATGTCCCCACCCTAAGCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC
 TTACACCTCTACTTGAGTATGTCCCTAACCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT
 GTCCCCCGTGTGCGCATGTGTTTCAAGGTGAGCCTCTCCAGAAGTGAGATCATGGACAAAAA
 GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCCAGCAAGAAGCTGAAC
 TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCT**TGA**AGTAACAAGTTTAAATGTTTCA
 GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT
 CCAAAAACACAAGTAGAAATTTCTAACAAATGAAATATATTACAGGCAGGTACCCACTAACCA
 AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC
 AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT
 AACAAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTTGGTCATGGTGGGCCAGCTGCAAAGCG
 TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC
 ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA
 ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTGAGGCTCTGGGCTCACCTCCATCTCCAGA
 GCATCCCCCTGCCTGCAGTTGTGGCAAGAACGCCAGCTCAGAATGAACACACCCACCAAGA
 GCCTCCTTGTTTATAACCACAGGTTACCCTACAAACCACTGTCCCCACAAACCCTGGGGAT
 GTTTTAAAACACACACCTCTAACGCATATCTTACAGTCACTGTGTCTTGCCTGAGGGTTGA
 ATTTTTTTTAAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

188/330

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLI FHIYLSMSPTLSPRSPQGW
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ
AQQEAEELTPRPAGVVPGA

188/330

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCAAAG
ATGGAGCTCTCGAAGGCCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG
TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCTTG
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACAACCTGGGAGACTGGGGATGA
CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAAACTGTGGAAG
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA
GCGGTTGATGGAGAAGGCTTCCCTCCCTCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA
TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTATCAGCTTCCTCCTGCT
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACCTGAGCGCCTTTGCTG
CTGTTTCCTCTGTCCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTACAAGTC
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA
ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCATAGTAAGAGCTTCAAGGAAAAC
CCGAAGTGCCTACCACATCACCATCAGTGTTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC
CGTGGGTCCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGGCCAGCCAGGAGCTG
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC
CTAAGGGATTCCCTGGGTGCCACTGCTCTCTTTTCCTCTACAGCTCCATCTTGTTTCACCCAC
CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC
TAAACCATGGAGATAAAAAGAAGAGTAAATACACTTCCCGACCTTAAGGATCTGAAA

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWVFGTQKVPKPLCEKGLAAKCFDMP
VSLDGDNTSTQEVVQYNWETGDDRFSSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR
GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLTTCLE
LWLGLLHGLALLHLLHGVGCHHLQHVVHVDGAGVQVQA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 191

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAGATGTGGTTACCCCTTGGTCTCCTG
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA
GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG
TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCCTCCTCCCAA
GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAG
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT
AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTTCAG
GACATTCGCCCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT
ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAA
GGGAAGCAACAGGAACCTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAACTCTTTATTACTTTGGG
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG
AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCAATGGTTTCTCCAATTATGCC
CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

192/330

FIGURE 192

MWLPLGLLSLCLSPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQSGME
HRNHLCFCDLYDRATSPPLKCSLL

192/330

194/330

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM
SNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFFLVFGSND
DFSWQQW

194/330

[illegible]

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNLSSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLLSLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVVIINEYGSCTCHQQPARECEV

FIGURE 197

GTGCAAGGAGCCGAGGCGAGATGGGCGTCCTGGGCCGGGTCCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGGCGCCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTTCCGCCTAGTGCCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCCGCAGCATCTCGGCTCTGGGCCGGACGTTACGCGCGACGAGGACCTGG
CTGTTTTCTGCGGTCCCGCGCGGGCCGCCTACGCTTCCACGGGCCGGGCGCGCTGAGCGTG
GGCCCCGAGGACTGCGCGGACCCGTCGGGCTGCGTCTGCGGCAACGCGGAGGCGCAGCCGTG
GATCTGCGCGGCCCTGCTCCAGCCCCCT

198/330

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE
GHAUSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA
PGLFFVDAERVPCRHDDVFFPPSASFVRVGLGPGASPVVRVRSISALGRTFTRDEDLAVFLASR
AGRLRFHGPALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

198/330

FIGURE 199

[illegible]

FIGURE 201

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCCTCCAGTCACCCCTCCCGCCGTACCCGCGGCGCGC
 CCGAGGGAGTCTCCTCCAGACCCTCCCTCCCGTTGCTCCAACTAATACGGACTGAACGGATCGCTGCGAGGGT
 GGGAGAGAAAATTAGGGGGGAGAAAGGACAGAGAGAGCAACTACCATCCATAGCCAGATAGATTATCTTACACTG
 AACTGATCAAGTACTTTGAAAATGACTTCGAAATTTATCTTGGTGTCTTCATACTTGCTGCACTGAGTCTTTC
 AACCACCTTTTCTCTCCAAGTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT
 TATATAAAGTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT
 TTTATTACAAAACCTACCCTAACCATTTACTTTGGTAACTGGCCTCTTTGCAGAGAATCATGGGATTGTTGC
 AAATGATATGTTTGATCCTATTCCGGAACAAATCTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTTT
 GGGAAGAAGCGACACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGACCCATGTGGCCCCGA
 ACAGATGTAAAAATACATAAGCGCTTTCCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG
 AGTTGCCAAAATTGTGAATGGTTTACGTCAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG
 ATGACATGGGCCACCATTGTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCAATTCAGATATTGACAAGAAGTTA
 GGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAAGTGATCATGG
 AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCACTATAACCTGATTG
 ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCT
 CATCCTAATCTTACTGTTTACAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTC
 ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC
 ACGGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTACGCCATGGTCCCTGCCTTCAGAAAGAATTTT
 TCAAAAGAAGCCATGAACCTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA
 CAATGGATCATCTGGAATGTCCAGGATCTGCTCAATTGAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA
 CTATACTCCTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTATACCCCTTATTTTATAGGGGTC
 TCTCTTGGCAGCATTATAGTGATTGTTATTTTTTTGTAATTTTTCATTAAGCATTTAATTCACAGTCAAATACCTGC
 CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAATGTTACTTTGAAGTGGATTGTCATA
 TTGAAGTGGAGATTCCATAATTATGTCACTGTTTAAAGGTTTCAAATTTCTGGGAAACCAGTTCCAAACATCTGC
 AGAAACCATTAAGCAGTTACATATTTAGGTATACACACACACACACACACATACACACACACGGACCAAA
 ATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGTCTCCATTGTTCACTGTAGCATAGGGATAGATAAG
 ATCCTGCTTTATTTGGACTTGGCSCAGATAATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATAT
 ATTGCACTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC
 TTGATTGAAAATGACAACCTTTTTGCACCCATGTACAGAATACTTGTTACGCATTGTTCAAACGAAGGAAATT
 TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTGTA
 AAATTAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATCCCAACAGCAGAATGCAACTGTGGGCAT
 TTCTTGCTTATTTCTTCCAGAGAACGTGGTTTTTCAATTTATTTTCCCTCAAAGAGAGTCAAATACTGACAG
 ATTCGTTCTAAATATATTGTTTCTGTATAAAATTATTGTGATTTCTGATGAGTCATATTACTGTGATTTTCA
 TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTTCAAGCAATGGCCTGAATAGAAGCAACCAGGCA
 CCATCTCAGCAATGTTTTCTTGTGTTGTAATTATTGCTCCTTTGAAAATTAAATCACTATTAATTACATTAA
 AAATCAAATTTGGATAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPINRKSFSLDHMNIYDSKFWEEATPIW
ITNQ RAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTLDQSPVAAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTTGTGATCCGCGATTTCGCTCCACGGGCGGGACCTTTGTAACCTGCGGGAGGCCAG
GACAGGCCACCCCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC
AGAGAGGCCAAGCCCCTTGCCTTGGGTACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA
CCAGATCCAGAGGCAACAGGGAC**ATG**GCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC
AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA
CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG
AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCTGACGTTGCC
CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTT
CAGCTCCACAGTTTTAGGTCATCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC
TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG
GTATTCACCTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT
ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCCGTCGTGGTGG
TGGTCTCATTTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC
CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT
TAAGACACGTTTCAAGACGGCAACTCTTAAGGTAAAACAGATGAATGTACAATTGGCCGCCA
AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCTGGAC**TGA**TGAGTTTGCTGTATC
AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT
CACACAGCCACCGTGAAAGTCCTGGAGTAAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG
CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTA
TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAGAGAACAAAACC
AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

204/330

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEEEQPPPTPV
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSHRFQVIIICLVVLDALLVLAELILDL
KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLLSSFTTSLRSWMPVVVVVSFILDI
VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLRLKQMNVQLAAKIQHLEFS
CSEKPLD

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
 LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILK
 SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW
 WRHLVAGGGAGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
 NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMAL
 RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS
 ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG
 LYRGLAPNFMKVIPAVSISYVYENLKITLGVQSR

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

208/330

FIGURE 208

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD
AGTYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVD
QGANFSEVSNTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTESEIKRR
SHLQLLNSKASLCVSSFFAISWALLPLSPYMLK

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 209

GAATTTGTAGAAGACAGCGGCGTTGCC**ATG**GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT
GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC
AGCGGGTTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG
AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC
CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACTCTCTGGCTATCATCTTCA
CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC
GAGTGCGGGACGCAGCTCTGTGGATCTCGACATACTGTGTTAGTTCCTTCCCAGAACCCAT
CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT
GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG
GTTTGGGAT**TGA**AAGACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT
GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT
GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC
AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT
GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT
CCTCCTCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG
AGTTTTGTAACCTTCAAGTGCTGTTCACTGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA
CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCTGCTAACATCTCAGGCTCCCAGCCCA
GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT
CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG
GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA
GTGGCAGATGCAGGAAATGAGCTGTCTATTAGCCTTGCCCTGCCCCACCCATGAGGTAGGCAG
AAATCCTCACTGCCAGCCCCCTTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCCACCTGAC
TCCAGCACACCTGGCGAGTAGTAGCTGTCAATAAATCTATGTAAACAGACAAAAAAAAAAAAA
AAA

211/330

FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG
GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTGCGAAAAGATTCCGCAATAAACT
TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT
TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCATGAAGGGCATCCTCG
TTGCTGGTATCACTGCAGTGCTTGTTCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT
TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG
CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT
TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT
GCTGAAGAACACTTTTCATTTTGTAAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG
CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG
AATCTAATGGAACCTCCTGTCTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTGTC
TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC
CAACGTGAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGAGGAGTCA
TCTTTCGAAAGTTTGAGTGTGCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACT
TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGCCCTTGCCAGCCTCCTTCTTCG
GGGACTGCTGCCCTGAGGTCCCTGGGGCTGCACTTTGCCAGCACCCCATTTCTGCTTCTCTG
AGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTCCCTGCTCTGCCCCGTTTAACTGC
CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA
AAGCACTGGTTCATTCACTGCCAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 213

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA
GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCC**ATG**GTCCCCGCCGCCG
GCGCGCTGCTGTGGGTCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA
ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC
ACGGTGTCCACCGGCTTTAGCCGGTCTGCCCATTAACGAGGAGGATGGGTCTTCAGAAGA
GGGGGTTGTGATTAATGCCGGAAGGATAGCACCAGCAGAGAGCTTCCAGTGCGACTCCCA
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG
ACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCAACCTCACCCACAG
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC
TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCT
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGCACCTATCAACAATGTC
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT
GCCTCTCAGAGCACCAACAGTACCAGGACCACCACTACCCCTTCCCCACCATCCACCTCAG
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCTGCCAGCCCTGGCTTTTTGGAAACGGGTCA
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA
GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACCTGGAAAA
CACAA

FIGURE 214

MVPAAGALLWVLLLNLGPRAGAQGLTQTPTMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAAELLAATVSTGFSSAINEDGSSEEGVVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLMWPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQCPCNRLREECPLDTSLCDTNCAQSTTSTRTTTTTFFPTIHLRSSPSLPPASPCPALA
FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPF SRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAFWSTVFHTRDSDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYSLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLQLGLSLLELLDFPPLFWVLDAAHAIWHISTIPVHVLFSSFLEDDSLYLL
KESEDKFKLD

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

FIGURE 217

[illegible]

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEEL
DAEVLEV FHP THEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGRDLINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLRTL VQEKGTEVLAVRVVTL LYDLVTEKMFAEEEAELTQEMSPEKLQQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

Important features:**Signal peptide:**

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCTTCCC
 CTTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCTCGTGGGGTGCCTTGCCACCCCA
 CGCGGACTCCCCAGCTGGCGCGCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCACCCCC
 TTCCACCTGACCAGCC**ATG**GGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGC
 CCGGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCGCTTCGCGTTATCATCCTGGT
 CGCAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG
 TCCATGTGACCGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCT
 GTCTCTGTCTTCTACAGGAGGTGTTCCGCTTTCCTACTACAAGCTGCTTAAGAAGGCAGA
 TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT
 ATGTTTCTGGTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT
 GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCTATTACTTCCTGACTTC
 AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG
 CCTGTGAGAGGAGACGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG
 GGAATGACATTTCCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCATCTATGCAGTCACTGT
 TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC
 TCTTGTGTAAGGAC**TGA**CTACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG
 CCCATGACTGAGCCCAGCCCCAGCCCGGGTCCATTGCCACATTCTCTGTCTCCTTCTCGTC
 GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCTTTTGTGACCGTTAGTCTCTAAGCTT
 TACCAGGAGCAGCCTGGGTTTACGCCAGTCAGTACTGGTGGGTTTGAATCTGCACTTATCCC
 CACCACCTGGGGACCCCTTGTGTGTCCAGGACTCCCCCTGTGTGAGTCTCTGCTCTCAC
 CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGGCAGGAGGACAGTCGGGTGAT
 GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC
 CTGGGGTGCCCTCCTGATGTCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG
 GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA
 GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC
 AGGCCTGAGGGGGAACCATTTTTTGGTGTGATAAATAACCCTAACTGCCTTTTTTTCTTTTTT
 GAGGTGGGGGGAGGGAGGAGGTATATTGGAACCTCTTCTAACCTCCTTGGGCTATATTTCTC
 TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT
 GGGGGAAGGAAGGAAGTGCATGTTTGGGAACCTGGCATTACTGGAATAATGGTTTTAACCT
 CCTTAACCACCAGCATCCCTCCTCTCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC
 CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG
 GAGATTTTTTTGTAGTTTTTAATTGGGGTGTGGGAGGGGCGGGGAGGTTTTCTATAAACTGT
 ATCATTTTCTGCTGAGGGTGGAGTGTCCATCCTTTTAATCAAGGTGATTGTGATTTTGA
 AATAAAAAAGAATTTGTAAAAA
 AA

FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIILVAGAFFWLVSLLLASVVWFILVHVTDR
SDARLQYGLLI FGA AVSVLLQE VFRFAYYKLLKKADEGLASLSE DGRSPISIRQMAYVSGLS
FGIISGVFSVINILADALGPGVVG I HGDS PYYFLTSAFLTA A ILLHTFWGVVFFDACERRR
YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

FIGURE 221

AAGCTGGTTTAAGGAAGCAGAGGAGGGTTAGATTCGTTGAGTGAGGACGGAAGATCAACCCA
TTTCCATTCCGCCAGATGGCCTATGTTTCTGGTCTCTCCCTTCGGNATCATCAGTGGTGTNT
TNTCTGTTATCAATATTTTGGCTGATGCANTTGGGCCAGGTGTGGTTGGGATCCATGGAGAC
TCACCCTATTANTTCCTGANTTCAGCCTTTNTGACAGCAGCCATTATCCTGCTC

FIGURE 222

GACCGACCGTTCAGATGCCCCGGTTCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTTTTNTCTGTTATCAATATTTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCATTATAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCCTTTCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTCCTGTCTGGTCAGGCCCCCACCCTTCCCACNTG
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTGCGGTTGCGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC
CGACCGGTCAGATGCCCCGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNTCGTGGGGTCGCGTTGCCACCCACGCGGACTCCC
CAGNTGGCGCGCCCCCTCCCATTTGCCTGTCTGGTCAGGCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTGGGGCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA
CCGACCGGTGAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCCAGGAGGAGGCAG
 TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC
 TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG
 TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCCTCCTTGCCCTGTCTGGAGGCTGCT
 AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCCGT
 CCTTGTGGTTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCC**ATG**GCTACAGCAAGACCCC
 CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT
 CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA
 GGACCTGGGAGCTGGGGCCGGGAAGACGCCCGGTGGATGACAGCAGCAGCCGCATCATCA
 ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC
 CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG
 GAAGAAAGTTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC
 AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCCTGGCTACTCCCACCCTGGCCACTCT
 AACGACCTCATGCTCATCAAACCTGAACAGAAGAATTCGTCCCCTAAAGATGTCAGACCCAT
 CAACGTCTCCTCTCATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA
 CCAAGAGCCCCCAAGTGCACCTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT
 CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA
 CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC
 TGCAGGGACTCGTGCTCCTGGGGAGATTACCCTTGTGCCCGGCCAACAGACCGGGTGTCTAC
 ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCAT
 CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG
 ACCCTCATTCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCT
 CCTGGACTCAGGGTCTGCTTCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG
 GAACAATTTCCAAAACCTGTCCAGGGCGGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCAT
 CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA
 CTGAGAAGTGGAATAAAAAA

FIGURE 226

MATARPPWMWVLCALITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS
PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRIRPTKDVRPINVSSHCP SAGTKCL
VSGWGTTKSPQVHFPKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGRDSCQGDSGGP
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 227

ATGGTCAACGACCGGTGGAAGACCATGGGCGGCGCTGCCCAACTTGAGGACCGGCCGCGCGA
 CAAGCCGCAGCGGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC
 TGCTGGCTGTAGCTGTACCGGTGCCGTGCTCTTCTGAACCACGCCACGCGCCGGGCACG
 GCGCCCCACCTGTCTGTCAGCACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGGA
 AAGGGCGGACAGCTCGCACCTCAGCATCCTCATTTGACCCGCGCTGCCCGACCTCACCGACA
 GCTTCGCACGCCTGGAGAGCGCCAGGCCTCGGTGCTGCAGGCGCTGACAGAGCACCAGGCC
 CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGCCGACCAGCTGCC
 CCGGCTGCTGGCCCGAGCCTCAGAGCTGCAGACGGAGTGCATGGGGCTGCGGAAGGGGCATG
 GCACGCTGGGCCAGGGCCTCAGCGCCCTGCAGAGTGAGCAGGGCCGCTCATCCAGCTTCTC
 TCTGAGAGCCAGGGCCACATGGCTCACCTGGTGAACCTCGTCAGCGACATCCTGGATGCCCT
 GCAGAGGGACCGGGGGCTGGGCGGGCCCCGCAACAAGGCCGACCTTCAGAGAGCGCCTGCCC
 GGGGAACCCGGCCCCGGGGCTGTGCCACTGGCTCCCGGCCCGGAGACTGTCTGGACGTCTC
 CTAAGCGGACAGCAGGACGATGGCGTCTACTCTGTCTTTCCCACTACCCGGCCGGCTT
 CCAGGTGTACTGTGACATGCGCACGGACGGCGGGCTGGACGGTGTTCAGCGCCGGGAGG
 ACGGCTCCGTGAACCTTCTTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCACC
 GGGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACAGGCTGCCTACGAGCT
 GCACGTGGACCTGGAGGACTTTGAGAAATGGCACGGCCTATGCCCGCTACGGGAGCTTCCGGC
 TGGGCTTGTCTCCGTGGACCTGAGGAAGACGGGTACCCGCTCACCCTGGCTGACTATTCC
 GGCACGTGCAGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTACACCAAGGACCGTGA
 CAGCGACCATTGAGAGAACAACTGTGCCGCTTCTACCGCGGTGCCTGGTGGTACCGCAACT
 GCCACACGTCCAACCTCAATGGGCAGTACCTGCGCGGTGCGCACGCCTCCTATGCCGACGGC
 GTGGAGTGGTCTCCTGGACCGGCTGGCAGTACTCACTCAAGTTCTCTGAGATGAAGATCCG
 GCCGGTCCGGGAGGACCGCT**TAGACT**GGTGCACCTTGTCTTGGCCCTGCTGGTCCCTGTGCG
 CCCATCCCCGACCCACCTCACTCTTTCTGTAATGTTCTCCACCCACCTGTGCCTGGCGGAC
 CCACTCTCCAGTAGGGAGGGGCCGGGCCATCCCTGACACGAAGCTCCCTGGGCCGGTGAAGT
 CACACATCGCCTTCTCGCCGTCCCCACCCCTCCATTTGGCAGCTCACTGATCTCTTGCCCTC
 TGCTGATGGGGCTGGCAAACCTGACGACCCCAACTCCTGCCTGCCCCACTGTGACTCCGG
 TGCTGTTTGGCGTCCCCTGGCCAGGATGGTGGAGTCTGCCCCAGGCACCTCTGCCCTGCCC
 GGCCAAATACCCGGCATTATGGGGACAGAGAGCAGGGGGCAGACAGACCCCTGGAGTCCCTC
 CTAGCAGATCGTGGGGAATGTCAGGTCTCTCTGAGGTGAGGTCTGAGGCCAGTATCCTCCAG
 CCTCCCAATGCCAACCCCCACCCCGTTTCCCTGGTGCCAGAGAACCCACCTCTCCCCCAA
 GGGCCTCAGCCTGGCTGTGGGCTGGGTGGCCCCATCCTACCAGGCCCTGAGGTGAGGATGGG
 GAGCTGCTGCCCTTGGGGACCCACGCTCCAAGGCTGAGACCAGTTCCCTGGAGGCCACCCAC
 CCTGTGCCCCGGCAGGCCTGGGGTCTGCAGTCTCTTACCTGCTGTGCCACCTGCTCTCTG
 TCTCAAATGAGGCCCAACCCATCCCCACCCAGCTCCCGGCCGTCTCTTACCTGGGGCAGC
 CGGGGCTGCCATCCCATTTCTCCTGCCTCTGGAAGGTGGGTGGGGCCCTGCACCGTGGGGCT
 GGACTGCGCTAATGGGAAGCTCTTGGTTTTCTGGGCTGGGGCTAGGCAGGGCTGGGATGAG
 GCTTGTAACAACCCCAACCAATTTCCAGGGACTCCAGGGTCTGAGGAGGCCAACCTTGCC
 GCCTTGGGGGTGATGACCCCTTCCCTGAGGTGGCTGTCTCCATGAGGAGGCCAACCTTGCC
 ATTGACCGTGGCCACCTGGACCCAGGCCAGGCCCGGCCGAGTGGTCAAGGGACAGGGGA
 CCACCTCACCGGGCAAATGGGGTGGGGGGACTGGGGCACCAGACCAGGCACCACCTGGACA
 CTTTCTTGTGAATCCTCCCAACACCCAGCACGCTGTCATCCCCACTCCTTGTGTGCACACA
 TGCAGAGGTGAGACCCGAGGCTCCAGGACCAGCAGCCACAAGGGCAGGGCTGGAGCCGGG
 TCCTCAGCTGTCTGCTCAGCAGCCCTGGACCCGCGTGCGTTACGTGAGGCCAGATGCAGGG
 CGGCTTTTCCAAGGCCTCCTGATGGGGGCTCCGAAAGGGCTGGAGTCAGCCTTGGGGAGCT
 GCCTAGCAGCCTCTCCTCGGGCAGGAGGGGAGGTGGCTTCTCCAAAGGACACCCGATGGCA
 GGTGCCTAGGGGGTGTGGGGTTCGTTCTCCCTTCCCCTCCCCTGAGGTTTGTGCTTAAAA
 AACAATAAATTTGACTTGGCACCCTGGGGGTGGTGGGAGAGGCGGTGTGACCTGGCTCTC
 TGTCCCAGTGCCACCAGGTCATCCACATGCCAG

FIGURE 228

MVNDRWKTMGGAAQLEDPRDPKQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELDLADQLPRLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRRDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAAYARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRSDHSENNCAAFYRGAWWYRNCHTSNLTNGQYLGAHASYADG
VEWSSWTGWQYSLKFSEMKIRPVREDR

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT
 TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT
 CACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG
 ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATCCAGAGCCC
 CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC
 TTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTACTACC
 AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC
 CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC
 TGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGATACAG
 AACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG
 GACTGTAAATATTTCTGCCTTAGTGAAAACCTCTACCATGCTGAAGATAAACAAACAAGAAGA
 CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT
 TGGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACCTCTGAACTG
 TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG
 GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA
 TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGACTGAATTCGCC
 CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT
 TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCCTCGTG
 TTTCTGTTTACAGGATCACCAGCATTTCTGAGCTTGGGTATATGCACGTATTTAACAGTCACA
 AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG
 GCTTAGAGATAACTTTTAGCTCTCTTTCTCTCAATGTCTAATATCACCTCCCTGTTTTTCAT
 GTCTTCCTTACACTTGGTGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC
 ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCAGATTGTACC
 AGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCAT
 CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACCTTGATATGAG
 AATCTCAAATCTCAATGCCTTATAAGCATTCTTCTGTGTCCATTAAGACTCTGATAATTG
 TCTCCCCTCCATAGGAATTTCTCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG
 AACTACCGTCCCCGATATTCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA
 TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG
 GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

FIGURE 231

AATTTTCACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGNTGG
ATGATGATGGGACACCACCATGAGCCTGCATTNTCAAGCTTTTGCCACAATTCGGCATCCAG
AGCCCCGGCGCACAGAGCACAGGGNTCCTTTTTTCAACGTGGCGACCAGTGGCCCTGACCCTG
CTGACTTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTA
CTACCAGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATA
CGTCCCAAGAGTTGCAATTTNTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTNTGCAGCAT
GTGGCTGAAAACTCTGTGCTGAGCTGTATAACAAAGCTGGAGGAACCTTGAAGGAGGGCAA
AGTNTCCTCATNTACTATACACACACCACTTCCC

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGGCTC
 CGGGGATTTCGGCTCGGGCCGCTGGCTCTGCTCTGCGGGGAGGGAGCGGGCCCCGCCGCGGGG
 CCCGAGCCCTCCGGATCCGCCCCCTCCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT
 CTGGGGGTTTCGCCGGGGCCGGGGACCCGCGGTCCGGGGCGCC**ATG**CGGGCATCGCTGCTGCTG
 TCGGTGCTGCGGCCCGCAGGGCCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCTGAGCCT
 GCTCAGCGTACCTGGGTGGAGGAGCCGTGCGGCCAGGCCCGCCCCAACCTGGAGACTCTG
 AGCTGCCGCCGCGCGGCAACACCAACGCGGCGCGCCGGGCCAACTCGGTGCAGCCCCGGAGCG
 GAGCGCGAGAAGCCCGGGGCCGGCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC
 CTACCACCCTGCACAGCCCCGGCCAGGCCGCCAAAAAGGCCGTAGGACCCGCTACATCAGCA
 CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC
 ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC
 GGGCGCACGGGGCCCGCGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC
 CCATTGGACACCTGCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGAC
 TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG
 CCACCTCAGCCTGGCCTCCGCCGCCACCTGTACCTGGGCCGGCCCCAGGACTTCATCGGCG
 GAGAGCCCCCCCCGGCCGCTACTGCCACGGAGGCTTTGGGGTGCTGCTGTGCGCGCTACTG
 CTGCAACAACCTGCGCCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCGCCCTGA
 CGAGTGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
 GGGTGCACCTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGCAGGAGGGGGACCCTCAT
 TTCCGAAGTGCCCTGACAGCCCCACCTGTGCGTGACCTGTGCACATGTACCAGCTGCACAA
 AGCTTTTCGCCCCAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA
 TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT
 CCAGCACCATCCCGCCCCGGCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA
 GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCACTGCGTGGGGCTGACCGGGCTG
 ATGTGGCCGATGTTCTGGGGCAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG
 CGGCTCCAGAAGCAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCCGGGTATGGA
 ATACACGCTGGACTTGCAGCTGGAGGCACTGACCCCCCAGGGAGGCGCGCGGCCCTCACTC
 GCCGAGTGCAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCTGTGCCCTATGTCACT
 GAGGCCCTCACGTCTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
 CTTCTTGAGAGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC
 TGCTGCTACTGTATGAGCCGCGCCAGGCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT
 GTCAAGGCCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG
 TGTGCAGACAGCCGCACCCTCACCCTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCTG
 TGGACACACTGTTCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC
 TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCCTTCTTTCCCATGCATTTCCAAGCCTTCCA
 CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCCAGAGCTGGGCCGTGACACTGGCCGCT
 TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACTCCGACTACGTGGCAGCCCGTGGG
 CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGGATGTGTACGAGCT
 GTTCCTCCACTTCTCCAGTCTGCATGTGCTGCGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT
 ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC
 GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG
 CAACAGCACCT**TGA**CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCCACTT
 CTCCCCAAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC
 AAGCTGGCCCCACTGGTCCCCTCTCTGGCTCTGTGGTCCCTGGGCTCTGGACAAGCACTGGG
 GGACGTGCCCCCAGAGCCACCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT
 GCTGATTTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC
 CTCTGGGCCCTGGGGGCTGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGGAGCTGAGGAGGGG
 GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA
 AGTGTGGAAAAA

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELP PRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLL
QHGDDEFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCIL DATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFRSALT AHPVRDPVHMYQLHKAFARAELE RTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGI PAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLTAL EELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGP GPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLES LDVYELFLHFSS LHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST

FIGURE 234

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT
TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT
AGCTATTAGCCAATTTCGGCAGGGCCCGCTTTTGTAGAAGCTTGATTTCCCTTTGAAGATGAAAG
ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACCTCGGGGCGATTGGCTGGGAA
CTGTATCCACCCAAATGTCACCGATTTCTTCCCTATGCAGGAAATGAGCAGACCCATCAATAA
GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGC AAAAG
AGGGTTGCTCAACGCCCCGCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG
GCGGAGGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT
TTCCCCGCCCCCTGAGACCCTGCAGCACCATCTGT**CATG**GCGGCTGGGCTGTTTGGTTTGAGC
GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCCGCGTCCGCTGGGA
ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC
CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTTGTATGAGAAGAACCCA
GACTCCCATGGTTATGACAAGGACCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT
CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA
GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC
CTTCCCATCATGGAATCCAACCTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAG**TG**
ACCAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCTGCCTGCCATTCTGAC
CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

235/330

FIGURE 235

MAAGLFGLSARLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE
DENLYEKNPD SHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER
LVKYREANGLPIMESNCFDPSKIQLPEDE

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 237

GCGGCGGCT**ATG**CCGCTTGCTCTGCTCGTCCTGTTGCTCCTGGGGCCCCGGCGGCTGGTGCCT
TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACCTTGTTCATCACCCCGCTGCCTTCCGGGG
ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTTCGGAGCTTCAGCGGGAAGGAGTG
TCCCATTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA
GCTGCACCTGTCAATCACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCACCCCTTCCTGC
AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA
TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGA
CTCCACCAACACAGTCACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG
ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC
CCCTGGAAGAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA
TCGCTTGTTCCACACCAGCTACCACTCCAGGCAGTGCATATCCGCCCTGTTTGCAGAAATG
CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC
ATCACGGGGCAGGGAAAGAAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCCTCACGGA
GCCCTGCCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACTACAACCAGGACA
ACGAGACATTAGAGGTGCACCCACCCCGACCACTACATATCAGGACGTCATCCTAGGCACT
CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACCTCTCGAAACCT
CAACATCCAGCTCAAGTGGAAGAGACCCCGAGAGAATGAGGCCCCCCCCAGTGCCCTTCCTGC
ATGCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC
AACACCCACCCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCG
GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAACCAAGTTACATCC
ACTACCAGCCTGCCCAGGACCGGCTGCAACCCACCTCCTGGAGATGCTGATTGAGCTGCCG
GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA
CACGCCAGATCCTAACCATGGCTTCTATGTCAGCCCATCTGTCTCAGCGCCCTTGTGCCCA
GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCCA
GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC
GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT
GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT
GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCACTCT**GA**TT
CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT
TCTGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTGAGGGC
CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTGTAATTTGAATTAA
CTTAGAAATTCATTTCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA
AGTGGTCCGTGGCTGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACACAGAAAGGTC
GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTGTAGTGGA
TGGAGTTTACTGTTTGTGGAATAAAAACGGCTGTTTCCGTGGAAAAAAAAAAAAA

FIGURE 238

MPLALLVLLLLGPGGWCLAEP PRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASF KPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSSISWELRQTLSVVFDAFITG
QGKKDWSLFRMF SRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRP PENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAKPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLLVLNLPDPDFSMYPNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

FIGURE 239

CAACATGGGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC
TGCTTCAAGTCCGATCCTCCCCAGTGTACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCCTGTGAAGGAACTGGAAGAAGGAGGAA
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC
TCCTCCTCTACCAGGTGTCCTCAGAAATGATGCTGGGTCCCTTTCTACCTCTGGGGGTCACTC
TCACTTGGCACCTGCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAAG
TCCAGACCCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAACTGGTCCTCACCAGATG
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAAGTGTAAAGAAATAAATAT
TTGCTGTTTATAATCCAA

240/330

FIGURE 240

MGSSSFVLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC
YLHCGFKCVIPVKELEEKGKDEEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 241

AAACCTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG
 AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC
 TCTAGAACCCGACCCACCACC**ATG**AGGTCCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG
 CGTCCAGTGGTTCCTTGCTTCTGGCTGTCCTGGTCTTCTTTCTCTTCGCCTTGCCCTCTTTTA
 TTAAGGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCT
 CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA
 TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCACACCA
 CCGGAGACAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCAC
 ACAGCACAGAGGGCAGCATGGAAGAGCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC
 ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC
 AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG
 GTGTGAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTTCCAAAAGTCA
 GCACAGAATGCTGGCTCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA
 CAGCAGTCATCCCACCTAAGGAGAAGAAACCTCAGGCCACCCACCCCTGCCCTTTCCAG
 AGCCCCACGACGCAGAGAAACCAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG
 GGATTTTGAGGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG
 TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAACTCTTTCTGCCCAACCTCACTCTC
 TTCCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC
 CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG
 TGCCCCAGCAGCAGCTGCTCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGT
 GCCGTGGTGGGCAACGGGGGCATCCTGAACAACCTCCACATGGGCCAGGAGATAGACAGTCA
 CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAGGCTACGAACAGGATGTGGGGACTC
 GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT
 CGGGGTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCTGGAAGGCAC
 CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT
 TCTGGTTTCAGGCACAGACCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG
 TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA
 TGGTGGCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC
 AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCATGAGCGCTTTTCTGAT
 CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA
 GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCTGGTC
 CCGGAACCTGCCAAAGCCAAGAAC**TGA**CCGGGGCCAGGGCTGCCATGGTCTCCTTGCTGCTC
 CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA
 GCTCCAAGCCCTTCAGGAGTTCCAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT
 GGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCTGAGGCCAGG
 GATTTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCCTGCTGAAAAACACTCTT
 CCAGTCCAAAAGCTTCTTGATACAGAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG
 GTTTGAATTCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACCTCACTAC
 AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCTTG
 TCTTTAAGCTATTTGACAACCTTACGTGTTGTAGAAAACCTGATAATAATACAAATGATTGTT
 GTCCATGGAAAGGCAAATAAATTTTCTACAGTGAAAAAAAAAAAAA

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPEKKPQATPPPAPFQSPPTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
NQSEWDRLEHFAPPPFGFMELNYSLVQKVVTFRPPVPQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRSLGALIKGYEQDVGTRTSFYGFATAFSLTQSLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTLGDAHWRIRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

244/330

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

FIGURE 245

GGGCTGGGCCCCGCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG
CCCGACCCCGGCGCGCCAGCCCCACC**ATG**CCACCCGCGGGGCTCCGCCGGGCGCGCCG
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTTCTGTTGCTACCTGT
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTTGAAGGCCAGGAGATTCCAATGACAGGCATC
CCAGTGACGCCAGTATACCCATAACCCCAAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCAATATCCACTCTACCCAGCTGGGCCCC
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC
TGAGGAACCAGCCATGTCTCTGCTGCCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA
GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA
ACTATGAGGGGTGGGGGGAGGGCTTGGAATTATGGGCTATTTTACTGGGGGCAAGGGAGG
GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT
TCCGTGAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG
CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLLVLGAPLVLAGEDCLWYLDRNGSWHPGFNCEFFTFC CGTCYHRYC
CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCF LCSCCYLYRRRQQLQSP
FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPP
YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 247

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGCGGCGCAAGGGTGAGGGCGGCCCCAGAA
 CCCCAGGTAGGTAGAGCAAGAAG**ATG**GGTGTCTTCTGCCCTCAAATGGTCCCTTGCAACCATG
 TCATTTCTACTTTCCTCACTGTTGGCTCTCTTAACCTGTGTCCACTCCTTCATGGTGTACAG
 CACTGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCTTGGAAATAAAATACGACTTC
 CTGAGTACGTCATCCCAGTTTATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACC
 TTCTGGGGAACACGAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCCTGCA
 TAGTCACCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGG
 AAGAACCCCTGCAGGTCTTGGAAACACCCCTCAGGAGCAAATTGCATGCTGGCTCCCGAG
 CCCCTCCTTGTCGGGCTCCCGTACACAGTTGTCACTATGCTGGCAATCTTTCGGAGAC
 TTTCCACGGATTTTACAAAAGCACCTACAGAACCAAGGAAGGGGAAGTGAAGGATACTAGCAT
 CAACACAATTTGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGATGAACCTGCCTTC
 AAAGCAAGTTTCTCAATCAAAATTAGAAGAGAGCCAAGGCACCTAGCCATCTCCAATATGCC
 ATTGGTGAAATCTGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGA
 AGATGAGCACCTATCTGGTGGCCTTTCATTTAGATTTTGAGTCTGTGAGCAAGATAAC
 AAGAGTGGAGTCAAGGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGC
 ACTGGATGCTGCGGTGACTCTTCTAGAATTTTATGAGGATTATTTTACGATACCGTATCCCC
 TACCCAAACAAGATCTTGCTGCTATTCCCGACTTTTCACTGCTGGTGCTATGGAAAAGTGGGGA
 CTGACAACATATAGAGAATCTGCTCTGTTGTTGATGCAGAAAAGTCTTCTGCATCAAGTAA
 GCTTGGCATCACAGTGACTGTGGCCCATGAACCTGGCCACCAGTGGTTTGGGAACCTGGTCA
 CTATGGAATGGTGGAAATGATCTTTGGCTAAATGAAGGATTTGCCAAATTTATGGAGTTTGTG
 TCTGTCACTGTGACCCATCCTGAACCTGAAAGTTGGAGATTATTTCTTTGGCAAATGTTTTGA
 CGCAATGGAGGTAGATGCTTTAAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTG
 CTCAGATCCGGGAGATGTTTGTATGATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATG
 CTAAGGGAGTATCTTAGCGCTGACGCATTTAAAGTGGTATTGTACAGTATCTCCAGAAGCA
 TAGCTATAAAAAATACAAAAAACGAGGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAG
 ATGGTGTAAAAGGGATGGATGGCTTTTGCTCTAGAAGTCAACATTCATCTTCATCCTCACAT
 TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGTTT
 TCCCTAATAACCATCACAGTGAGGGGGAGGAATGTACACATGAAGCAAGAGCACTACATGA
 AGGGCTCTGACGGCGCCCCGGACACTGGGTACCTGTGGCATGTTCCATTGACATTCATCACC
 AGCAAATCCAACATGGTCCATCGATTTTGTCTAAAAACAAAAACAGATGTGCTCATCCTCCC
 AGAAGAGGTGGAATGGATCAAATTTAATGTGGGCATGAATGGCTATTACATTGTGCATTACG
 AGGATGATGGATGGGACTCTTTGACTGGCCTTTTAAAGGAACACACACAGCAGTCAGCAGT
 AATGATCGGGCAAGTCTCATTAAACATGATTTTCAAGTCTGTCAGCATTGGGAAGCTGTCCAT
 TGAAAAGGCCTTGGAATTTATCCCTGTACTTGAAACATGAAACTGAAATTATGCCCGTGTTC
 AAGGTTTGAATGAGCTGATTCCCTATGTATAAGTTAATGGAGAAAAGAGATATGAATGAAGTG
 GAAACTCAATTCAGGCCTTCCTCATCAGGCTGCTAAGGGACCTCATTGATAAGCAGACATG
 GACAGACGAGGGCTCAGTCTCAGAGCAAATGCTGCGGAGTGAACCTACTACTCCTCGCCTGTG
 TGCACAACTATCAGCCGTGCGTACAGAGGCGAGAAGGCTATTTTCAAGAAAGTGAAGGAATCC
 AATGGAAACTTGAGCCTGCCTGTGACGTGACCTTGGCAGTGTGTTGCTGTGGGGGCCAGAG
 CACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTTTTTCTTTGTCCAGTACTGAGAAAA
 GCCAAATTGAATTTGCCCTCTGCAGAACCCAAAAATAAGGAAAAGCTTCAATGGCTACTAGAT
 GAAAGCTTTAAGGGAGATAAAATAAAAACTCAGGAGTTTCCACAAATTTCTTACACTCATTGG
 CAGGAACCCAGTAGGATACCCACTGGCCTGGCAATTTCTGAGGAAAAACTGGAACAACTTG
 TACAAAAGTTTGAACCTGGCTCATCTTCCATAGCCACATGGTAATGGGTACAACAAATCAA
 TTCTCCACAAGAACACGGCTTGAAGAGGTAAAAGGATTCTTCAGCTCTTTGAAAGAAAATGG
 TTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATGAAAGAAACATCGGTTGGATGG
 ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAAAAA**
 TTCTCCCTTGCCCGGTTCTGTTATCTCTAATCACCAACATTTTGTGAGTGTATTTTCAA
 ACTAGAGATGGCTGTTTTGGCTCCAACCTGGAGATACTTTTTTCCCTTCAACTCATTTTTTGA
 CTATCCCTGTGAAAAGAATAGCTGTAGTTTTTCATGAATGGGCTTTTTTCATGAATGGGCTA
 TCGCTACCATGTGTTTTGTTTCATCACAGGTGTTGCCCTGCAACGTAAACCCAAGTGTGGGT
 TCCCTGCCACAGAAGAATAAAGTACCTTATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 248

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH
 YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLOVLE
 HPPQEQIALLAPEPLLVLGPLYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA
 ARMAFPCFDEPAFKASFSEIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA
 FIISDFESVSKITKSGVKVSVYAVDPKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA
 IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL
 WLNEGFAKFMEFVSVSVTHPELKVG DYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD
 DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG
 FCSRSQHSSSSSSHHWQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD
 TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL
 TGLLKGTHTAVSSNDRASLINNAFQLV SIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP
 MYKLMKCRDMNEVETQFKAFILRLRLDLIDKQTTWDEGSVSEQMLRSELLLLACVHNYQPCV
 QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSSTEKSQIEFALC
 RTQNKEKLQWLLDESFKGDKIKTQEFQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS
 SSIAHMVMGTTNQFSTRTRLEE VKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIR
 VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

FIGURE 249

CAGCCACAGACGGGTCATGAGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT
TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATGGGCCCCGTGGGTATGACTGAGA
ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGG
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCGTGCCAGGAGACCGGCAGTGTC
CTACCTGTGTGCAGCCCCCTTGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG
GGCGCCACTCATTGTTATGATGGGTACATTATCTCTCAGGAGGTGGGCTGTCCACCAAAT
GAGCATTACAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTGAACCACACCAGACAAATCG
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGGTTTGCCCTTCCTGCTTAACTCTATTACCCCCACGATTCTTCACCGCTGCTGA
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTATATCTACTCACCTAACAGCA
AACTGGGGAGAGCCTGGAGCATCCGGAATGCCCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRGGLIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESITWGVGLALAPALWWGVVC
PSC

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRGGLIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESITWGVGLALAPALWWGVVC
PSC

FIGURE 251

GCGACGGGCAGGACGCCCCGTTTCGCCTAGCGCGTGCTCAGGAGTTGGTGTCCTGCCTGCGCT
CAGG**ATG**AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG
CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG
CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCCGACGGCCTGGAAGAGTCG
GCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTCGT CAT
GGAAAAATTGGTCCCATTTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC
TGGTCCTAATGGAGAACCAGGCCTCCCATGTGTAGTGCAGCCAGCTGCGCAAGGCCATCGGGG
AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTGCGC
GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA
CGCCCAGCTGTCTTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA
ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCCGTGTCTTCATCGGCATCAAC
GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA
GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGAGATGGTGGCCT
CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG
GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTTGGGGGCCCCACATGTCCCTGCAGGGTT
GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG
GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG
AAAGTGTTCTTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA
ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC
TATACAATAAAATCTTTAAGTAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 252

MRGNLALVGVLISLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG
PTGEKGDMDKGQKGSVGRHGKIGPIGSKGEKGDSDIGPPGPNGEPGLPCECSQLRKAIGE
MDNQVSQLTSELKFINKAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN
GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS
GGWNDVACHTTMYFMCEFDKENM

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCTAGCCAGTTCTTGATCCTGCCAGACCACC
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACC**ATG**AGGATCATGCTGCTATTCACAGCCAT
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTCTTACCTTCAGTGAGGGTTCCTCGGCCC
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC
TTTAT**TAA**GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA
TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT
CCTACATTAAAAATATAATGTCTCTCTCTATTCTCAACAATAAAGGATTTTTGCATATGAA
AA

FIGURE 254

MRIMLLFTAILAFSLAQSGAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK
ALSQASTDPKESTSPEKRDMDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK
SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCCGCTGTGCCTGCTGTGCC
 CGCGCTGTGCGCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG
 GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC
 CCATAGTTGCTGCAGGAGTGGAGCC**ATG**AGCTGCGTCCTGGGTGGTGTATCCCCCTTGGGGC
 TGCTGTTCTGCTGCGGATCCCAAGGCTACCTCCTGCCAACGTCACTCTCTTAGAGGAG
 CTGCTCAGCAAATACCAGCACAAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA
 GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT
 CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC
 CTGGGCCACCAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCCTGTGATGG
 CTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTTGTTTGTGTTT
 TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG
 AAACCT**TAG**ACTCCCGGGGTTAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACCTACAG
 GCATGCACCATGGTGCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT
 TGCCCAGGCTGGTCTTGAACTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG
 CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA
 ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC
 TCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGGCCAGGG
 CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC
 CTTAGCCCGTGTGAGCCTCACTTTCCACTTGGAGAGTCCTTCCTCGCGTGGTTGCCATGACT
 GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG
 CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCCTCCCGGT
 GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG
 TGGAAACTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG
 GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT
 TTTAAACAAACCCACGTCCCAGCCTGGGTAAACATGGTAAAGCCCCGTCTCTACAAAAAATC
 CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG
 GTGGAGGTGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC
 AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

FIGURE 256

MSCVLGGV IPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNE SHSRVRRRAIPREDKEEILML
HNKLRGQVQPQASNMEY MVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR
GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

FIGURE 257

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACT**ATG**GG
GTCTGGGCTGCCCCCTTGTCCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGG
GTATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGC
TTCCTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCAC
CCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACA**TGA**CAGCCATTGAAGCCTG
TGTCTTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTT
CAGCAGGCCCCCACCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

258/330

FIGURE 258

MSGGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTS
VTLHHARSQHHVVCNT

FIGURE 259

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAAACAATTAACTGTTAGGATTGCAGT
TATGATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCTCCACTGTGTTCTGTGTGC
TATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTTGAAAATGAATGCTTAGTGATCTG
TGCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTTCCCCTCCC
CTCCGATTGTTCTAAAT**TAA**TTGAAAGATGTCTGCTGTGGAAAAGGCATGTATTTAAATCTG
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCTTGAAAGCCAATGGAAATACTTTTTTT
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA
GTAAAATAGAAACCTGTGTTTATTCTCAGGTATTTTAGAAACAACAGCCATCATTTTATTTT
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCATTTTCTC
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAAATTGATATTAATAAAACCAGA
ATAAAAGTTCATATCTACCC

260/330

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLP
SDCSK

Important features:

Signal peptide:

amino acids 1-29

260/330

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGAGCACCACCGAGGCCCTTGAGACATCCTT
GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTGCAGGATGATGGTGGCCCTT
CGAGGAGCTTCTGCATTGCTGGTTCTGTTCTTGCAGCTTTTTCTGCCCCCGCCGCAGTGTAC
CCAGGACCCAGCCATGGTGCTATTACATCTACCAGCGCTTTCGAGTCTTGAGCAAGGGCTGG
AAAAATGTACCCAAGCAACGAGGGGCATACATTCAAGAATTCCAAGAGTTCTCAAAAAATATA
TCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACTT
GGCACTGAGAGTTGAACGTGCCCAACGGGAGATTGACTACATAACAATACCTTCGAGAGGCTG
ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA
GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC
TTTGAANAATAGTGAAGAAGATGATGGACACACATGGCTCTTGATGAAAGATGCTGTCTATA
ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC
ATACGGGCATTTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC
CTGGCAGGGAAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTTCAATAACCAAGCAACTT
CTAATGAGATAATCAAATATAACCTGCAGAAGAGGACTGTGAAGATCGAATGCTGCTCCCA
GGAGGGGTAGGCCGAGCATTGGTTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT
GGATGAGCATGGGCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA
CAAAGATTGAGCCGGGCACACTGGGAGTGAGCATTTCATGGGATACCCCATGCAGAAGCCAG
GATGCTGAAGCCTCATTCCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA
GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGC
CCAACTTGTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCCAGAGAT
AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTTACAACTCCAGACAAAGAGAAA
GCTGCCTCTGAAGTAATGCATTACAGCTGTGAGAAAGAGCACTGTGGCTTTGGCAGCTGTTT
TACAGGACAGTGAGGCTATAGCCCCTTCACAATATAGTATCCCTCTAATCACACACAGGAAG
AGTGTGTAGAAGTGGAATACGTATGCCTCCTTTCCCAAATGTCACCTGCCTTAGGTATCTTC
CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAAA
CCTCCTGGCTCTCAAGGATGACCACATTCGATACAGCCTACTTCAAGCCTTTTGTTTTACT
GCTCCCCAGCATTTTACTGTAACTCTGCCATCTTCCCTCCCACAATTAGAGTTGTATGCCAGC
CCCTAATATTACCACCTGGCTTTTTCTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT
CAAATGTCTATTGATATTCTCCCATTTTCACTGCCCAACTAAAATACTATTAATATTTCTTT
CTTTTCTTTTCTTTTTTTTTGAGACAAGGTCTCACTATGTTGCCCAGGCTGGTCTCAAACCTCC
AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC
CACACCTGGCTTAAATACTATTTCTTATTGAGGTTTAACTCTATTTCCCCTAGCCCTGTC
CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAATATTAACATTTGAATATCGCTTT
CCAGGTGTGGAGTGTGTGCACATCATTGAATTCTCGTTTTCACTTTTGTGAAACATGCACAAG
TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC
TAGAAAATACTACAAATCCCATAGTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT
TGTTCACTACTCTTATAGTCAATCGTTTCATCGTTTTCAGCCTAAAAATAATAGTCTGTCCC
TTTAGCCAGTTTTTCTAGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC
AGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCCTCCTCTTGTCTTGCTGTCTCTGT
TTCTCTCTTTCTGCTTTAAATTCAATAAAAGTGACACTGAGCAAAAAAAAAAAAAAA

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL
QEAEKKIRTLNLASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNPKVYLLIGSRNNTV
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGLGVEHSWDT
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

119



FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQRIRKEKPQQHNFTHRLLAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTAAVSPCGRFVASCG
FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSLGALKK

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG
CAGTGTTTTGCCTTCACCCCAAGTGACCATGAGAGGTGCCACGCGAGTCTCAATCATGCTCC
TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT
GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT
GGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA
AGCACCACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC
CGTGCTCCATGGACTTGAAGAACATCAATTTTTTAGGCGCTTGCTGGTCTCAGGATACCCA
CCATCCTTTTCCTGAGCACAGCCTGGATTTTTATTCTGCCATGAAACCCAGCTCCCATGAC
TCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTGTCTAGTACGCACATATGCACACAG
GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG
CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTCCCTGCTCAGGCTGCCAGAGAGGTGGTA
AATGGCAGAAAGGACATTCCCCCTCCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG
CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCTGGGCACAGGCTCTTGGGT
GCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG
ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC
TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA
GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA
CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT
TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA
TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACCTCTTTCCA
AGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG
TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTGAGACCAGGGAGG
CAACTACACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAA

266/330

FIGURE 266

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCCAISLWLRGLRMCTPLGREGEECHP
GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 267

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAAATGTCTTTC
CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC
TGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG
CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA
GCAAAGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT
CCTCTGTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG
TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT
TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTCCACAAAGCGGAAGAT
GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCTTCCGAGCCTGGAACGGAG
GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCGTGGTGGGATCA
GGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT
TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAATGAT
TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTTCATGGGATGTATT
GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA
TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA
CAAAAATCTGAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA
TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG
TGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC
TCTACTAAAAATACAAAATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCG
GGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA
CACCACTGTATTCCAGCCTGGGTGACTGAGACTCTAACTAA

FIGURE 268

MSFLQDPSFFTGMWSIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKEPR
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEFGLGGVFV
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 269

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCG
GGCCAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA
AGCCCCCTTCCTCGGCGCTGCCAACCCGCCACCCAGCCC**ATG**GGCGAACCCCGGGCTGGGGCTG
CTTCTGGCGCTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACA
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTG
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG
CACCTACCGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG
ACTCCAAGGAGACGGTGCAGGGCTGCCTGCCCATC**TAG**GTCCCCTCTCCTGCATCTGTCTCC
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGG
GCTATTCACCTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

270/330

FIGURE 270

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIV
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIV
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

FIGURE 271

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT
TTGGGATTTTAATTTTCAAACACAGCAGAATGACATTTTTCTGTCACTATTATTATTGTTG
GTATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCT
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC
CTACCAAAGCTGTCAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA
GGGTTAATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATG
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATA
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT
TTATTAATTTTTTAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGA
TCATATAATTTGATACAAATAAAAGAAAAGTGTTCTCTCCCCTTACAGAAATTGACATTTTAA
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

FIGURE 273

GCCAGGAATAACTAGAGAGGAACA**ATG**GGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTTCT
 GTGCCTGCTGCACCAAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA
 TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAATAATTGAACAAATAGAGGAT
 ATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGATTTTTTTTTCAAAAA
 TGTATCTATATTAATTCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG
 AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA
 TACACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCT
 TCTACTTGGAaaaaaaacaaaatgaatatggaccaccaggcaaaactgTTTTGCCATGAGTGGG
 CTCACCTCCGGTGGGGAGTGTTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG
 TCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA
 GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACCTGTATG
 GAAAAGATTGTCAATTCCTTCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG
 CAAAGTATTGATTCTGTTGTTGAATTTTGTAAAGGAAAAAACCCATAATCAAGAAGCTCCAAG
 CCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAATTCTGAGGATT
 TTAaaaaacaccatacccatggTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC
 AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT
 AAATCGAATGAATCAAGCAGCAAAACATTTCTGCTGCAGACTGTTGAAAAATGGATCCTGGG
 TGGGGATGGTTCACCTTTGATAGTACTGCCACTATTGTAATAAGCTAATCCAAATAAAAAGC
 AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTTCCAT
 CTGCTCTGGAATTAAATATGCATTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT
 CCGAAGTACTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGATTGATGAAGTG
 AAACAAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT
 AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG
 GCCTCATTGATGCTTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAACTCCCTT
 CAGCTCGAAAGTAAGGGATTAACTACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT
 TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA
 TTTCTCTCTGGGATCCCAGTGGAAACAATAATGGAAAAATTCACAGTGGATGCAACTTCCAAA
 ATGGCCTATCTCAGTATTCAGGAACCTGCAAGGTGGGCACTTGGGCATACAATCTTCAAGC
 CAAAGCGAACCAGAAACATTAACCTATTACAGTAACCTTCTCGAGCAGCAAAATCTTCTGTGC
 CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATT
 GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCAT
 TGAATCACAGAATGGACATACAGAAGTTTTGGAACTTTTGGATAATGGTGCAGGCGCTGATT
 CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT
 AGCTTAAAAGTTTCGGGCTCATGGAGGAGCAAAACACTGCCAGGCTAAAATTACGGCCTCCACT
 GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA
 GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTAGCCGAACAGCATCCGGA
 GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCCTGACCAATACCCACCAAGTCA
 AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG
 GAGATAATTTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGA
 GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAAGAAATGCAACCC
 ACATATTTATTGCCATTAAGTATAGATAAAAGCAATTTGACATCAAAAGTATCCAACATT
 GCACAAGTAACTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC
 TACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT
 TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATT**TGA**ACCTTA
 ACGAAGAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTAAAAAACAAAACAAATGTAAGT
 AAAGGATATTTCTGAATCTTAAATTCATCCCATGTGTGATCATAACTCATAAAAATAATT
 TTAAGATGTGCGAAAAGGATACTTTGATTAAATAAAAAACACTCATGGATATGTAAAAACTGT
 CAAGATTAAAAATTAATAGTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAAC
 AAAGATCCTTTTTCTACTGATACCTGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT
 GATATTTCAAATTCATCAAGAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAA
 GGAGAGCAAATAAACAAACATTTGGAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
 AA

FIGURE 274

MGLFRGFVFLLLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTASTY
LFEATEKRFFFFKNVSILIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVVE
FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLV
LDKSGSMGGKDRLNRMNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDQSEVLLLTGDGNTASSCIDEVKQSGAIVH
FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSTIMENFTVDATSKMAYLSIPG
TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNQDVYSRYFTAYTENGRYSLKVRAGH
GANTARLKLRPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRTASGGAFVVSQV
PSLPLPDQYPPSQITDLATVHEDKIIILTWTAPGDNFQVGVQRYIIRISASILDLRDSFDD
ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIQVTLFIP
QANPDDIDPTPTPTPTPTPTDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 275

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG
 GGCAGGGGTGACAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCTATTTCAAGGAAAG
 ACGCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACC
 CCCAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTT
 GTGTTGGAGGAGAGAACCCTTTGTGGGGCTGCGTCTCTTAGCAGTGCTCAGAAGTGAAGTGCCTGA
 GGGTGGACCAGAAGAAAGGAAAGGTCCCTCTTGCTGTTGGCTGCACATCAGGAAGGCTGTGATGGG
 AATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAAA
 GTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATG
 CAAGCAGCTCCGGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGG
 GCCCCGGCTTTGAGGGATGCCACCGGTTCTGAGCGCATGGCTGATTCTGAAATGATGATGGTTCGCC
 GGGGGCTGCTTGGTGGATTTCCCGGGTGGTGGTTTGTCTGGTGTCTCCTCTGCTGTGCTATCTCTGT
 CCTGTACATGTTGGCCTGCACCCCAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGCCAACAGC
 CCCACGGGGAAAGGAGGGGTACCAGGCCGTCTTCAGGAGTGGGAGGAGCAGCACCAGCAACTACGTGA
 GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGCAGGAGAGGAGTGC
 TGGGCAGTACCAAGCCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAG
 GCCGACCTCCTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG
 CCACAGAGTATGCAGCAGTGCCTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGTGGAGACTGG
 CCTTACCCGCCAGCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTGTGGAAAGCCATT
 GAATCAGCCTTGGAGACCCCTGAACAATCCTGCAGAGAACAGCCCCAATCACCGTCTTACACGGCCT
 CTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTTCAA
 AGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTACAGCCCCATCATGAAAGTG
 AAAAATGAAAAGCTCAACATGGCCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGG
 ACAAGTTCCGGCAGTTTCATGCAGAAATTTAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCT
 CACTGTTGTTTACTTTGGGAAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAA
 GCTGCCAACTTCAGGAACCTTACCTTCATCCAGCTGAATGGAGAATTTTCTCGGGAAAGGGACTTG
 ATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGCTCTTCTCTTTTTTCTGTGATGTGGACATCTACTT
 CACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTTTATCCAGTT
 CTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAGC
 AGCTGGTCATAAAGAAGGAAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTCAATATCG
 GTCAGACTTCATCAATATAGGTGGGTTTGTGAGACATCAAAGGCTGGGGCGGAGAGGATGTGCAC
 CTTTATCGCAAGTATCTCCACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACC
 TCTGGCATGAGAAGCGCTGCATGGACGAGCTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAA
 GGCCATGAACGAGGCATCCACCGCCAGCTGGGCATGCTGGTGTTCAGGCACGAGATAGAGGCTCAC
 CTTTCGCAAAACAGAAACAGAAAGCAAGTAGCAAAAACATGAAGTCCCAGAGAAGGATTGTGGGAGA
 CACTTTTTTCTTTCTTTTGAATTTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAGGACG
 ACAAAGAATTTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTCTGTTGGGCTTTTTAC
 AACAGAAATCAAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACA
 AAGGCAGAAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGTTTTGTATGGTGTTTACAATACT
 GAGACCTGTTGTTTTGTGTGCTCATTGAAATATTCATGATTTAAGAGCAGTTTGTAAAAAATTCAT
 TAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTCTAGG
 AATGCTAAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACTAGTGAGTACATTAAAGTA
 AAATAAAATGGACCAGAAAAGAAAAGAAACCATAAATATCGTGTCTATATTTCCCCAAGATTAAACCA
 AAAATAATCTGCTTATCTTTTTGGTGTCTCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAAT
 GCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGA
 GCACAAGTTGGCCTACATTTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTCA
 GTTCAAAGCATCAAATTTGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAAT
 GTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAA
 GAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAA
 AAGGAAACTCATTCAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACATTTTCTCCTCA
 GAAGTAGGGACCGCTTTCTTACCTGTTTAAATAAACCAAAGTATACCGTGTGAACCAACAATCTCT
 TTTCAAAACAGGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATAT
 ATATATATATATTGTGAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTTTGTCTACAT
 GTTATCCACCCAGGCCAGGTGGAAGTAACCTGAATTTATTTTAAATTAAGCAGTTCTACTCAATCA
 CCAAGATGCTTCTGAAAATTCATTTTATTACCATTTCAACTATTTTTTAAAAATAAATACAGTTA
 ACATAGAGTGGTTTTCTTCATTTCATGTGAAAATTTATAGCCAGCACCAGATGCATGAGCTAATTTATCT
 CTTTGAGTCTTGTCTGTTTGTCTGCTACAGTAACTCATTTGTTTAAAGCTTCAAGAACATTTCAAGC
 TGTGGTGTGTTAAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTTAATTAACAC
 AGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTTGTGGATATGAA

FIGURE 276

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPV RKDKRDELVEAIES
ALET LNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP
MKVKNEKLN MANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCR
LNTQPGKKVFYPVLF SQYNPGIIYGHDAVPPLEQQQLVIKKETGFWRDFGFGMTCQYRSDFI
NIGGFDDL DIKGWGGEDVHLYRKYLHSNLIVVRTPV RGLFHLWHEKRCMDELTP EQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

FIGURE 277

GAAAGAATGTTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC
AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT
ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA
GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT
ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC
AATCAGCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAAAT
CTGGAATTTTTTAAAAATCCCTTCCACACTTGACACCACCCATGGACCCATCTGTGCCCATCTG
GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT
CAGGGATCTGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT
AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCCTCTGATCCCCCTGGACATGAAGGG
GGGCATATTAATGATGCCTTCATGACAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT
TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA
CCAAGAGCAGATCATATATTTTGTTCACCATTCCTTTTGTAATAAATTTTGAATGTGCT
TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC
TCAAAATATTCTAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG
TAGTTATTGATTTAAGCATTTTTTAGAAATAAGATCAGGCATATGTATATATTTTACACTTC
AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT
TGAAAATGGATCCTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG
TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGAATT
TTATCCTGTTATCACACCAACAGTTGATTATATATTTTCTGAATATCAGCCCCAATAGGAC
AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG
TAATAATCATCTCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 278

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVAFSMRKVP
NREATEISHVLLCNVTQRVSFVVTDP SKNHTLP AVEVQSAIRMNKNRINNAFFLNDQTL E
FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC
ENMITIENGIPSDPLDMKGGILMMPS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 279

AACTCAAACCTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG
 GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATGT**
 ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG
 CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC
 TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCCGTGGGTGATGCTCTAACAGTGACCTGGA
 ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC
 CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA
 TGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATAACCTGCCAGGTGA
 AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA
 CGTTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT
 AATAGTAATTGTAGTGGTCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC
 ATAAAGTGGTGGAGATAAAATCAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT
 GTTTATTTAGAAGACACAGACT**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA
 GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT
 TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC
 AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTTCA
 GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAACAAGAAGCTACATTTTTTGCCCTTAA
 GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC
 AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA
 ATGTGTTTACTCTCTTTCTTCCACATTCTCAATTAAAGGTGAGCTAAGCCTCCTCGGTG
 TTTCTGATTAAACAGTAAATCCTAAATCAAAGTGTAAATGACATTTTTTATTTTTATGTCTC
 TCCTTAACTATGAGACACATCTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATT
 TTTGTCTG

FIGURE 280

MYGKSSTRAVLLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSEFAPVGDALTVT
WNFRPLDGGPEQFVFYYHIDPFQPMSEGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ
VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIVIVVVLFQHYRKKRWAER
AHKVVEIKSKEEERLNQEKVSVYLEDTD

FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGAATGC
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTGAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGCG
GCAAACCTCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAGTAAACCACAGGCTGG
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTACCACATTAGCAATTAAACTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG
GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 285

GTC**ATG**CCAGTGCCTGCTCTGTGCCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA
GCGGCCCCCATGGGCGGCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG
GACCCTGCAGCTGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTCAGC
CGGGGCCGGGATGCAGCCAGGAACCTCGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCC**TGA**ATCTGCCTGGATGGAAGTGA
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG
CCTGTTCACTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA
CCCTTTCATGCCTACACACCCTCATTAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 287

GGCAACATGGCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT
 CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAGCCAGGA
 AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC
 TGGACAGAAGTCAATGCCTTGAAGGAAATTCAAGCCCTGCAGACAGTCTGTCTCCGAGGCAC
 TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG
 AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC
 CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA
 CATGGTCACGGAAGGCAAGTTTGTGACGTCAACGGAATCGCTATCTCCTTCCTCAACTGGG
 ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACCTGTGTCTCTTCTCCCAATCAGCTCAG
 GGCAAGTGGAGTGATGAGGCCTGTGCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC
 TAAATTAGGTCTTTCTCCAATGTGTCTCCAAGCAAGATTCATCATAACTTATAGGTTTCATGA
 TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT
 GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT
 TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT
 AAACAGACTAAAATCTTCTCTCTAGTCTTCTCACTTGTACAAACCCAGTTTGTTTTCAA
 AAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTTAGGCTACCTGAAAGATT
 TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCTACATCAGAGACTCTAGGT
 GCTATATAATCCAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC
 TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT
 GCCATATCAGAACACAAACCCCTGAAGAGGTTCTGATTTGATTTTTTTTTTTTTCTTCATGCC
 TACCCTTTTTTTTGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT
 CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT
 CAAAGAAGCAGATTGCATGATAAACGGAAATAGAAAAAAGAACCTACATTTATTTTGCTTT
 AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCTTTTTTTACATTT
 TCGTATATTTATTTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT
 TGGAAGCTGAAAACCTGAATTTAAAGAATGCTATCTTGAAAATTGCATACGTCTGTGCAATT
 TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT
 TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTTAGAAGCAAACAATTTT
 AAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT
 CCCACTTTGCAAACCTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTTCATTGCTCA
 ATAATAAAGCCTGAATTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLLDQTTSHTSRLKARKHKKRRVRDKDGDLDKTQIEKLWT
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSSQAQ GK
WSDEACRSSKRYICEFTIPK

MAQQACPRAMAKNGLVICILVITLLLLDQTTSHTSRLKARKHKKRRVRDKDGDLDKTQIEKLWT
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSSQAQ GK
WSDEACRSSKRYICEFTIPK

FIGURE 289

GCGAGGACCGGGTATAAGAAGCCTCGTGGCCTTGCCCGGGCAGCCGCAGGTTCCCCGCGCGC
CCCGAGCCCCCGCGCCATGAAAGCTCGCCGCCCTCCTGGGGCTCTGCGTGGCCCTGTCCTGCA
GCTCCGCTGCTGCTTTTCTTAGTGGGCTCGGCCAAGCCTGTGGCCCAGCCTGTCGCTGCGCTG
GAGTCGGCGGCGGAGGCCGGGGCCGGGACCCTGGCCAACCCCTCGGCACCCTCAACCCGCT
GAAGCTCCTGCTGAGCAGCCTGGGCATCCCCGTGAACCACCTCATAGAGGGCTCCCAGAAGT
GTGTGGCTGAGCTGGGTCCCCAGGCCGTGGGGGCCGTGAAGGCCCTGAAGGCCCTGCTGGGG
GCCCTGACAGTGTGGTTGAGCCGAGACTGGAGCATCTACACCTGAGGACAAGACGCTGCC
CACCCGCGAGGGCTGAAAACCCCGCCGCGGGGAGGACCGTCCATCCCCTTCCCCCGGCCCT
CTCAATAAACGTGGTTAAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAA

290/330

FIGURE 290

MKLAALLGLCVALSCSSAAFLVGS AKPVAQPVAALESAAEAGAGTLANPLGT LNPLKLLLS
SLGIPVNHLIEGSQKCV AELGPQAVGAVKALKALLGALT VFG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 291

TGAAGGACTTTTCCAGGACCCAAGGCCACACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT
 CCTTGGCCTCCGCAGCCGATCACATGAAAGTGGTGCCAAGTCTCCTGCTCTCCGTCTCCTG
 GCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCAGCCCC
 TCAGAACCCAGACCAGCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGG
 CCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGGCAGCAGCTT
 GCCAAGGAGACTTCAAACCTTCGGATTTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG
 CAACATGGTCTTCTCTCCATTTGGCATGTCCTTGGCCATGACAGGCTTGATGCTGGGGGCCA
 CAGGGCCGACTGAAACCCAGATCAAGAGAGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAG
 CCCGGGCTCCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACCT
 GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT
 TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCCTGCCTATGAATTTTCGCAATGCCTCA
 CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTCCCAAACCT
 GTTTGATGAGATTAATCCTGAAACCAAATTAATTTCTGTGGATTACATCTTGTTCAAAGGGA
 AATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCACCTGGACAAGTAC
 AAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA
 TTTTCGTTGTCATGTCTCAAACCTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCTCTCA
 TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA
 TGGCTCAGAAACATGAAAACCAGAAACATGGAAGTTTTCTTTCCGAAGTTCAAGCTAGATCA
 GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCCTTTG
 CTGACCTTAGTGAACCTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGA
 ACAGTGATTGAAGTTGATGAAAGGGGCACTGAGGCAGTGGCAGGAATCTTGTGAGAAATTAC
 TGCTTATTCCATGCCTCCTGTCATCAAAGTGGACCGGCCATTTCAATTCATGATCTATGAAG
 AAACCTCTGGAATGCTTCTGTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTATTAAATTCAGG
 ACATGCATAAGCACTTCGTGCTGTAGTAGATGCTGAATCTGAGGTATCAAACACACACAGGA
 TACCAGCAATGGATGGCAGGGGAGAGTGTTCCCTTTTGTTCTTAACTAGTTTAGGGTGTTC
 AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGC
 CTGAAACGGTGGACAGTGCTGAACCTTATATATATTTTTTCTTACACATACATACCTATGAT
 AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA
 TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACCTGATTATAGAGAAGGCTA
 CTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCA
 CATCCTGGGTGGGACAGAGCAGGACGATGCAAGATTCCATCCCCTACTCAGAATGGCATGC
 TGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTTCATTTAATGTTTTTGGACCATGGT
 TGACCATGGTTAACTGAGACTGCAGAAAGCAAACCATGGATAAGGGAGGACTACTACAAAA
 GCATTAAATTGATACATATTTTTTAAAAA

FIGURE 292

MKVVP S L L L S V L L A Q V W L V P G L A P S P Q S P E T P A P Q N Q T S R V V Q A P R E E E E D E Q E A S E E K A G E
E E K A W L M A S R Q Q L A K E T S N F G F S L L R K I S M R H D G N M V F S P F G M S L A M T G L M L G A T G P T E T Q I
K R G L H L Q A L K P T K P G L L P S L F K G L R E T L S R N L E L G L S Q G S F A F I H K D F D V K E T F F N L S K R Y F
D T E C V P M N F R N A S Q A K R L M N H Y I N K E T R G K I P K L F D E I N P E T K L I L V D Y I L F K G K W L T P F D P
V F T E V D T F H L D K Y K T I K V P M M Y G A G K F A S T F D K N F R C H V L K L P Y Q G N A T M L V V L M E K M G D H L
A L E D Y L T T D L V E T W L R N M K T R N M E V F F P K F K L D Q K Y E M H E L L R Q M G I R R I F S P F A D L S E L S A
T G R N L Q V S R V L R R T V I E V D E R G T E A V A G I L S E I T A Y S M P P V I K V D R P F H F M I Y E E T S G M L L F
L G R V V N P T L L

292/330

FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL
FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE
RPRLWVMPNHQVLLGPEEDQDHIYHPQ

FIGURE 297

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC
CACGAGGCTGCCGCATCCTGCCCTCGGAACAATGGGACTCGGCGCGCGAGGTGCTTGGGCCG
CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCC
ATGGCGGCATCTGCAAACATAGAGAATTCTGGGGCTTCCACACAACCTCCAGTGCTAACTCAAC
AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAACTTCCAACAGTACTGTGAAAC
CACCAACTTCAGTTGCCTCAGACTCCAGTAATAACAACGGTCACCACCATGAAACCTACAGCG
GCATCTAATAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC
TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG
TAACCCACAATAGTTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT
TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC
GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC
GGTATCGAACCATAGATGAACATGATGCCATCATTTAAGGAAATCCATGGACCAAGGATGGA
ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAACAATATTCT
CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA
AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT
TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTTGGCT
GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA
TGCCATCTGGGCATACAAATAAGAAGTTTGTACAGCACTCAGGATTTTGGGTATCTTTTGT
AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA
CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG
TGCTTTAAACTGTAGTAGTTGGTCTAGAAACAAATACTCC

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG
TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCAGCG**ATG**GCGACCCTGTGGGGAGGC
CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTTCGTGCCTGGCGCTTTCCTGTGCTGCTGCTGGC
GCAGCTGTCAGACGCCGCCAAGAATTTGAGGATGTCAGATGTAAATGTATCTGCCCTCCCT
ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT
CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA
ATGCAAATATGAAGAAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA
TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG
CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT
TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG
AATATGCACAGCAGCGCTGGAAGCTTCAAGTCCAAGAGCAGCGAAAGTCTGTCTTTGACCGG
CATGTTGTCCTCAGC**TAA**TTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA
CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTCATTTTAATACCTTGTTGATTTCACCAACT
GTTGCTGGAAGATTCAAACTGGAAGCAAAACTTGCTTGATTTTTTTTTCTTGTTAACGTA
ATAATAGAGACATTTTTAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG
TGACTTTTACTAATAAAAAATAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA
AGCACTCTCTTTTTTACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTG
TTGTTGTTGTTTTTTGTTTGTGTTTGGTGGGAGAGGGGAGGGATGCCTGGGAAGTGTT
AACAACTTTTTTCAAGTCACTTTACTAAACAACTTTTGTAATAGACCTTACCTTCTATTT
TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG
ACTTTTGCAGTCACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT
CTAAAATGCCTGGTGGCTTTTCACAAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG
CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG
GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGACACT
TGCAATAAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATATACACATTTG
TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT
AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT
TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTGAAACATCAA

FIGURE 300

MATLWGGLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENS
GHIYNKNIS
QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERS
SVTIKVTII IYLSILGLLLLLYMVYLT
LVEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDV
LARSRSRANVLNKVEYAQQRWKLQVQEQ
RKSVFDRHVLS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 301

GCACCTGCGACCACCGTGAGCAGTC**ATG**GCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTT
CTGGGCTTGTCTGGCTCTGTCGCTGCTGCTGCCAAGGCCTTCCTGTCCCGCGGGAAGCGG
CAGGAGCCGCCGCCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCA
CCAGGCACCCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGG
CATTTGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTG
ATGGGGCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTTATATATACTGTACATTCTATT
TAAGGTAAGTAGAATCATCCTAATCATATTACATCAAT**TGA**AAATCTAATATGGCGATAAAAA
TCATTGTCTACATTAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTA
AATCCTGCCTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTAGTTTTACATAAGAATG
TTTACTCAATGTTTAAGTGTTTTGCCCCAAAATTCACAATAACAAGGCAGAACTAGGACTT
GAACATGGATCTTTTGGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

302/330

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQT
PGARFQRSHLAEAFKAKGSGGGAGGGGSGRGLMGQIIPIYGFGI FLYILYILFKVSRIILI
ILHQ

1000 9000 8000 7000 6000 5000 4000 3000 2000 1000 0

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT
GGATTTGAAAGTTGAGAGCAGC**ATG**TTTTGCCCCACTGAAACTCATCCTGCTGCCAGTGTTAC
TGGATTATTCCCTTGGGCCTGAATGACTTGAATGTTTTCCCCGCTGAGCTAACAGTCCATGTG
GGTGATTCAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT
AGACTGGACTCTGTACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA
ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC
AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA
AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG
AGGAGCCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTCAGATGGGATGTGTTTTCCAG
AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA
GGAGATTGTATTTTCGTTACTACCACAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG
GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG
CTTCAAGGAGTGAGGGAGTCAGATGGAGGAACTACACCTGCAGTATCCACCTAGGGAACCT
GGTGTTCAAGAAAACCATTTGTGCTGCATGTCAGCCCCGAAGAGCCTCGAACACTGGTGACCC
CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTTGTGGGAATTGTC
TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA
GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG
AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA
CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA
CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACCTCACTTGAAAAAAGTCAGGTGGGG
GAATGCCAAAAACACAGCAAGCCTTT**TG**AGAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG
TGGAGACTCTCTCCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTC
CCAGCTGTCCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG
CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC
CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCGTT
GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGATGAGTTACTGGGAAGAATCAGAGATA
AAAACCAACCCAAATCAA

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWGHFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIH LGNLVFKKTIV
LHVSPEEPRTLVT PAALRPLVLGGNQLV IIVGIVCATILLLPVLILIVKKT CGNKSSVNSTV
LVKNTKKTNP EIKEKPCHFERCEGEKHIYSPI IIVREVIEEEE PSEKSEATYMTMHPVWPSLR
SDRNN SLEKKSGGMPKTQQAF

1000 900 800 700 600 500 400 300 200 100 0

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG
GTTCTACCCCTACTAAAGACAGGAAGATCATAACTGACAGATACTGAAATTGTAAGAGTTGG
AAACTACATTTTGTCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA
GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG
GCCCTGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG
ATGGTTGTCTGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA
TGAGAATGAAAATCGCACAGGAACCTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG
TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAAAC
TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACCTAACATGGGAAGAGAG
TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG
AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT
GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGAAGA
TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG
AGAACAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT
TAATGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT
ATGAATGCATCAGTAGCTGAAAAAAAAAAAAA

FIGURE 306

MQDEEDGYITLNIKTRKPALVSVGPASSSSWVRVMALILLILCVGMVVGLVALGIWSVMQRNYL
QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEEFL
EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 307

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCCG
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGGCGAAGTCGCGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCGAGCAAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGCTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
GTAGCCAGTGAGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGG
CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCKNKLNVFSRVKLFSGSKKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTDKEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA
MNSEGYLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMSHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG
 ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG
 CAACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCT
 CCATTTAGAGTGTAGCAAAGGAAAAAACACCAAGGTTGGGTCCTTCCTGACATTGGCAGTG
 CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC
 AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA
 ACCACGTCTTGGAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTG
 GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAACTCGAGGGTGGGATCCACTGAGG
 AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT
 GTGGAGGGGGGTACGTGAGGGGGGGGTCTGGGGCTTATCCTCAGGTCCTGTGGGTGGGGCAG
 CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG
 CGCGCTCCGGGCGCCTGCCGGTTTGGGGGTGTCTCCTCCCGGGGCGCTATGGCGGCGCTGGC
 CAGTAGCCTGATCCGGCAGAAGCGGGAGGTCCGCGAGCCCCGGGGGAGCCGGCCGGTGTGG
 CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG
 CTGTCCAAGGTGCGACTGTGCGGGGGGCGGCCCGCGCGGCCGGACCGCGGCCCGGAGCCTCA
 GCTCAAAGGCATCGTCACCAAACTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG
 ACGGAAGCATCCAGGGCACCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT
 GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC
 TGAGGGACTGCTCTACAGTTCGCCGCATTTACAGCTGAGTGTGCTTTAAGGAGTGTGTCT
 TTGAGAATTACTACGTCTGTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGCC
 TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTGATGAAGGGAAACCGAGTTAAGAAGACCAA
 GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC
 ACAGTGTCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCCTTGAAATGTAGTCCCTGGACTG
 GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACAACCTGT

FIGURE 310

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR
GPEPQLKGIVTKLFCRQGFYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY
MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

FIGURE 311

ATGGCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACG
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCT
TCAACCTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTTACCCCTGAATGCAAGTT
TAAAGAATCTGTTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGG
AATCTGGTAGAGCCTGGTTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA
GCACAAGTGCGTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDRPSASRRRSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLR
RQDPQLKGIVTRLYCRQGYLQMHDPD GALDGTKDDSTNSTLFNLI PVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELF TPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETV PKPGVTPSKSTSASAIMNGGKPVNKS KTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 313

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAAT
 GAAGGATGCAGGACGCAGCTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGA
 ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG
 AAATAAACAGAGTTAGACCCGCGGGGGTTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCC
 CTCCCCACCCCCAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAGTATGTTCATTT
 TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAA
 AGAACTGGTGTGGTGGTGTTCCTTTCTTTTTGAATTTCCACAAGAGGAGAGGAAATTAATAATACATCTGC
 AAAGAAATTTTACAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCA
 CAGTTGGATTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT
 TTTTTAAATTTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTCTTAACCACCTGGATTTCCATCT
 GGATGTTGCTGTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATG
 TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAAACAGGGCCCTATTTGACCCCT
 GCTTGTGGTGTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGT
 GCTCCTGCAGCAACCAGTTGAGCAAGGTGATTTGTGTTTCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC
 ACCAACACACGGCTGCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAG
 GCACTTGGAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA
 ACCTCAACACTCTGGAACCTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA
 CTGAAGGAGCTCTGGTTGCGAAACAACCCATTGAAAGCATCCCTTCTTATGCTTTTAAACAGAATTCCTTCTTT
 GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAAC
 TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAATCCCTAACCTCACACCGCTCATAAACTAGATGAG
 CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTCCAGGGTTTGATGCACCTTCAAAAAC
 GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGAGATCA
 ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGATCATCTAGAGCGGATACAT
 TTACATCACAAACCTTGGAACGTAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC
 GAACACAGCTTGTTGTGCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA
 ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT
 GAGCTGAAATGTGGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA
 TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTACAAATGTAAGTGTGCAAGATA
 CAGGCATGTACACATGTATGGTGAGTAATCCGTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCA
 GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG
 GACCACAGATAACAATGTGGGTCCCACTCCAGTGGTGCAGTGGGAGACCACCAATGTGACCACCTCTCTCACAC
 CACAGAGCACAAGGTGACAGAGAAAACCTTCACCATCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT
 GATGAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT
 GGTCAATTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCACGCCCCAACAGGACTGTTGAAATTA
 TTAATGTGGATGATGAGATTACGGGAGACACACCCATGGAAAGCCACCTGCCATGCCTGCTATCGAGCATGAG
 CACCTAAATCACTATAACTCATACAAATCTCCCTTCAACCACACAACAACAGTTAACACAATAAATTCATACAC
 CAGTTCAGTGATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATCTAAACA
 TTTACAGAGTTACAAAAACAAACAATCAAAAAAAGACAGTTTATTAAAAATGACACAAATGACTGGGCTAA
 ATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAACAAGAAAGAAATTTATTTATTAATAAATCTATTG
 TGATCTAAAGCAGACAAAAA

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
 VRKNLREVPDGI STNTRLLNLHENQIQI I KVN S F KHLRHLEILQLSRNHIRTIEIGAFNGLA
 NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
 YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDSLGNHLSAIRPGSFQGLMHLQKL
 WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNCDIL
 WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
 LKCRASLTLSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVDGMYTCMVSNVSGN
 TTASATLNVTAATTTTFFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNTTSLTPQ
 STRSTEKTFTIPVTDINSIGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQN
 HHAPTRTVEIINVDDEITGDTPMESHLMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS
 VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419,
 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
 391-397, 422-428, 433-439, 531-537

FIGURE 315

GCGCCGGGAGCCCATCTGCCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCCGGGCACAT
 GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCCAGCTCGCCCCGAGGTCCGTCCGA
 GCGCCCCGGCCGCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCCGCTCCGGGGATC
 GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA
 CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACCTTTGCCCTGCCACCATCAACTGGGGC
 TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
 GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
 AGTGGCCTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC
 CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT
 GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
 AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
 ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
 GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
 CCAGTGACACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACGTGTACAGT
 ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
 ATTTTCTCCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
 GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
 CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT
 AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
 CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
 CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
 ACGGTCT**TGA**ATTACAATGGACTTGACTCCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC
 TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
 GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCTTATAACAATACCAAACAAGCAAA
 AGGATGTAAGCTGATTTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
 AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
 AGGTGAATATACCTAAAACCTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTACAAATT
 TTCAAGAGGAAATGGGATGCTGTTTGTAATTTTCTATGCATTTCTGCAAACCTTATTGGATT
 ATTAGTTATTTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
 TGAGCTAACCACCTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
 TTCATTTGTCTATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
 AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAATAAC
 TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
 ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT
 CAGATGCCTCTAAGGACTTTCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
 TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
 CCAAACATAACCATTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
 GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
 TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
 ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

[illegible]

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFAFARSLSITTPPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHLVVLVKPSGARCIVDGSEEIGSDFKIKCEPKESLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSCTYTVRNRVGSQCLLRNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247,
245-251, 296-302

Myelin P0 protein.

amino acids 96-125

FIGURE 319

TGAAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGA
 ATGGTGAAGGTGCCTGTCTAACTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCC
 CTAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGC
 CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG
 AGAGATGAATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC
 CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG
 CTTCTCTGGAGCCAGGTATCAGGGGGCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT
 GAAGGGGGTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG
 CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT
 GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACGTGTTTTCAAAAA
 CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAAC
 TTGTTCTCATCGTGTCAAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC
 AGTGCACACAGGCGGTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC
 TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC
 TCTTGAATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCCTGTGTCATTTCA
 AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATT
 TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG
 GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACCTCTATTT
 AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG
 CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG
 TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTTGGCCTGTCTTTGGATT
 GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACGAACTTCAGAGCATGAAAATCACACT
 GTCTTCTGATATCTGCAGGGACAGAGCATTGGGGTGGGGGTAAGGTGCATCTGTTTGAAAAG
 TAAACGATAAAATGTGGATTAAAGTGCCAGCACAAAGCAGATCCTCAATAAACATTTTCATT
 TCCCACCCACACTCGCCAGCTCACCCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTTT
 TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC
 ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA
 CTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFCPPLLATASQMOMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNH
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL
TKALGEVDILLTWMQKFKYL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

FIGURE 321

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGCATGAAGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTC
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACCTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
TCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
GTGGGAGACAGCCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGACAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGC AAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETQLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHNDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG
AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACTGCACCTC
GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACGATGTTGGGGGCCCCGCCT
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCCATCAGAC
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA
TGAGCAGAAGATACCTCTGCATGGATTTTCAGAGGCAACATTTTTGGATCACACTATTTTCGAC
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC
CACCCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACCTTCAACACC
CCCATACCACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCCTGAACGT
GCTGAAGCCCCGGGCCCCGATGACCCCGGCCCCGGCCTCCTGTTTCACAGGAGCTCCCGAGCG
CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTGAGTGAAC
ACGCACGCTGGGGGAACGGGCCCCGAAGGCTGCCGCCCTTCGCCAAGTTCATCTAGGGTCTG
CTGG

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSWGGLIHLYTATARNSYHLQIHKNGHVD
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY
DVIHSPQYHFLVSLGRAKRAFLPGMNPPPYSQLSRNEIPLIHFNTPIPRRHTRSAEDDSE
RDPLNVLKPRARMTAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTAGGTGPEGCRPFA
KFI

Important features of the protein:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 325

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**
GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGTGCTGGGAGGTTGGGGTCTCTGG
GAGCTCTGCAGGCCCCAGCACCCGCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAACTCAAACGCTGAGCGCTGAG
ACCTCTTCTAGGGCCTCAACCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC
TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCCTTTGCACCGA
TGACAGCTCTGAAGAGGGCAAAGACACTCACAAATGGACATATTGACATTGGCTCACACCTCCA
CAGAAGCTAAGGGCCTGTCTCAGAGAGCAGTGCCCTCTTCCGACGGCCCCCATCCAGTCATC
ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT
GGTCCCCGGGATCTGATGTCACCTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG
GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACCTTCCAGCATCCCTGGGGCCTCAGA
CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCCTGTCC
ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCCACTCCCCACTAA
CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTGAGCAGTGGGCAAAACAACCTTC
CTTTGCTGGGAGCTCTGCTTCCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCCT
CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
GATCACAACTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCACGACTGCCCGGAC
GAGGCCGACCACAGACG**TGAG**TGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG
TGGCTTCCCCGGAAGACCTCACTGACCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
CGGGAACCTCCACGCCACGCGCCTCACTTCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA
ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCCTAGCCTG
GGCCCCCACCAGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG
AAGGGCAGCATGTCCAAGCCCCTAACCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC
CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCAGAGGTGTCCTTGGACTCACCTTGG
CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATACCCATCTGTGTGCTTCCATCCTGCA
TTAAATTTCACTCAGTGTGGCCCCAAAAAAA

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTL
 ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
 TTVQTTITGSDPEEAI FDTLCTDDSS EEAKTLTMDILT LAHTSTEAKGLSSESSASSDGPHPV
 ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTL LAEALVTVTNI
 EVINCSITEIETTTSSIPGASDIDLIPT EGVKASSTSDPPALPDSTEAKPHITEVTASAETL
 STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTL SGALVTVSRNPLEETSALS VETPSY
 VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
 PLPSVPPTTTNSSRGTNSTLAKITTS AKTTMKPQQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,
 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,
 404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-
 320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 327

GCGGAGCATCCGCTGCGGTCCTCGCCGAGACCCCGCGCGGATTTCGCCGGTCCTTCCCGCGG
 GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCTCTCGACGCCA
 GAGAGAAATCTCATCATCTGTGCAGCCTTCTTAAAGCAAACCTAAGACCAGAGGGAGGATTAT
 CCTTGACCTTTGAAGACCAAACTAACTGAAATTTAAATGTTCTTCGGGGGAGAAGGGAG
 CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT
 GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC
 AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAACTGCATTAATTCTTGCTGTTCAAC
 AAAAAACATATCAGGGGACAAAGCATGTAACCTTGATGATCTTCGACACTCGAAAAACAGCTA
 GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA
 AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG
 CCAAGAGTTACCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC
 TAGCCCATCATCACACAGATTATTCAAAGCCACCGATATCTCATGGAGAGACACACTTTCT
 CAGAAGTTTGGATCCTCAGATCACCTGGAGAACTATTTAAGATGGATGAAGCAAGTGCCCA
 GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTTACAATTTTCTCTGATCAAGAAA
 TAGCTCATCTGCTGCCTGAAAATGTGAGTGCGCTCCAGCTACGGTGGCAGTTGCTTCTCCA
 CATAACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC
 TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC
 AGCCTCCCACGACCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG
 GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA
 AACCATAACGTTTACAGAAATCTCCAACCTTAACCTTGAACACAGGGAATGTGTATAACCTA
 CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAACTGCTTCCTGGGAAGGT
 AGGGAGGCCAGTCCAGGCAGTTCTTCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC
 ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCTCTGTTCTGGTGATAGGCC
 TCGTCTCCTGGGTAGAATCCTTTCGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT
 TATTTGATCAATGGGATCTATGTGGACATCTAAGGATGGAACCTCGGTGTCTCTTAATTCATT
 TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG
 TATTTTGAAGACAGGAAAATGCCCCCTTCTGCTTTCTTTTTTTTTTTGGAGACAGAGTCTT
 GCTCTGTTGCCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC
 CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA
 CCACACCTGGGTGATTTTTGTATTTTATAGTAGAGACGGGGTTTACCATGTTGGTCAGGCTG
 GTCTCAAACCTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG
 CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTGAGAAGGAATGAAGTG
 GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAATATTAGCTAAAAACAAAGCTCT
 ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAATTCAACTGGCTTTTATGCAAA
 GAAACAGGTTAGGACATCTAGGTTCCAATTCATTACATTCTTGTTCCAGATAAAATCAAC
 TGTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTTATATGGATTCTTTTAAACTTATT
 CCAGATGTAGTTCCTTCCAATTAATATTTGAATAAATCTTTTGTACTCAA

328/330

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLKGI RGN EPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPN CYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQSFSSDQEIAHLLPENVSALPATVAVASPH TTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFEKWLLIGSLLFGVLFVLVIGLVLLGRILSESLRRKRYSR LDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

FIGURE 329

CTCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
 CCCAGGTTATGAAGCCCTGGAGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
 CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
 ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
 TCTTTACTGATCTCTCTGTTCTGTTTCCAGGACCCTGCTGTCTCCCTCCCCTTCTCCCAC
 CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC
 CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
 GCTGAGGCCCCCTCCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC
 TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGTGCTGAGCCTTCTGTGAGC
 CGCAGGCCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
 CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
 CCTTCCCAGGCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCAGAAAGCTTTCACCTCAGCCTCAGAG
 TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCAGGCTCTCCTCTTGATGTTCCA
 GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTTGGGC
 CTCATGCCAGTGTCGGACCCTGCCTTCCCTCCCACTCCAGACCCCACCTTGTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTTCAGGAAGCCT
 GTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
 GACTCTGAATTCTAACAATGCCAGTGAAGTGTGCACTTGAGTTTGAGGGCCAGTGGGCCCTG
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
 TCCAGGCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
 TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCCTTGGAAAAAATGATGAAGA
 AAACCTTGGCTCCTTCCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTCTGGTGGCTAGAGA
 GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
 ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGGGGGTGGTGGTAAAGTA
 GCACAACTACTATTTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
 GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCCTGGGTTCAGTGATT
 CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCTGAC
 CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
 TCTGGCCCTATTTCTTTTAAAAAGTGAAATTAAGAGTTGTTTCAGTATGCAAAACTTGGAAAG
 ATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTACCCATAGTCTCACCAGAGACTATCAT
 TATTTTCGTTTTGTTGTACTTCCCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTTCTT
 TTTACAGAGCAATTATCTTGTATATACACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
 ATCACTTTATTCAGCACTTCTCTGTGTTTTACAGACCTTTTTTATAAATAAAATGTTTCATCA
 GCTGCATAAAAAAAAAAAAAA

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPTEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSIRDSRQELSLIVTLWNLTQLQDAGEYWCGVEKRGPD ELLISLFV
FPGPCCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSI PMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE EKEAPSQAPEGD
VISM PPLHTSEEE LGFSKFVSA

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128